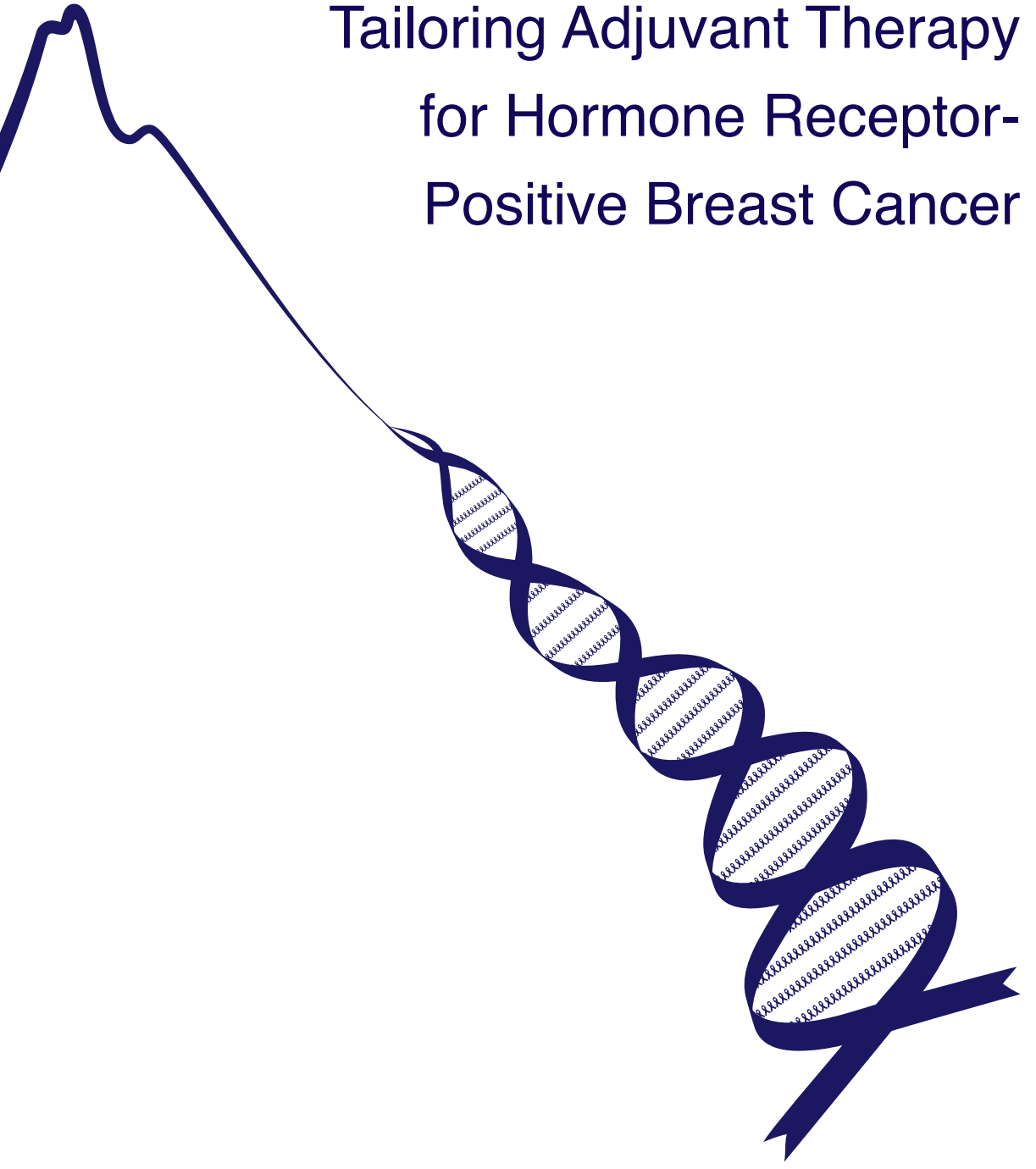
Tailoring Adjuvant Therapy for Hormone Receptor-Positive Breast Cancer



Erik J. Blok

Tailoring adjuvant therapy for hormone receptor-positive early breast cancer

Erik J. Blok

Lay-out: Ferdinand van Nispen, *my*-thesis.nl
Cover design: Dorijn Otto
Printed by: GVO drukkers & vormgevers

Tailoring adjuvant therapy for hormone receptor-positive early breast cancer
Thesis, Leiden University Medical Center, 2018

ISBN: 978-94-6332-334-5

Printing of this thesis was financially supported by: afdeling Medische Oncologie (LUMC), Agendia, Chipsoft, Pfizer and Philips Research

Tailoring adjuvant therapy for hormone receptor-positive early breast cancer

Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 31 mei 2018 klokke 13:45 uur.

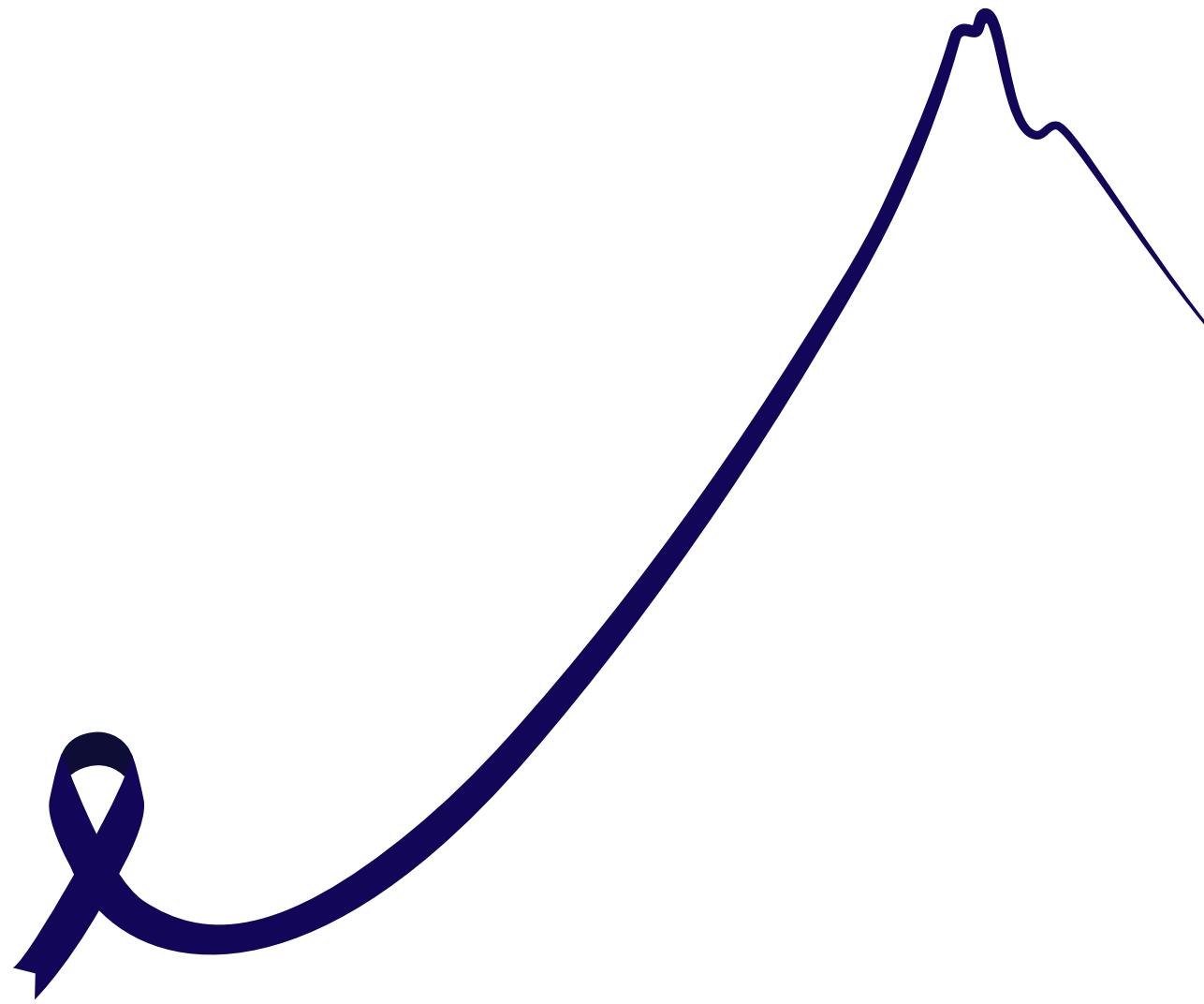
door
Erik Jan Blok
Geboren te Papendrecht
in 1989

Promotor:	Prof. dr. C.J.H. van de Velde
Copromotores:	Dr. J.R. Kroep Dr. P.J.K. Kuppen
Promotiecommissie:	Prof. dr. V.T.H.B.M. Smit Prof. dr. S.C. Linn (NKI/AvL) Prof. dr. ir. J.J.M. van der Hoeven (Radboud UMC)

Pap, deze is voor jou

Leo Blok
27/8/1959 – 22/10/2015

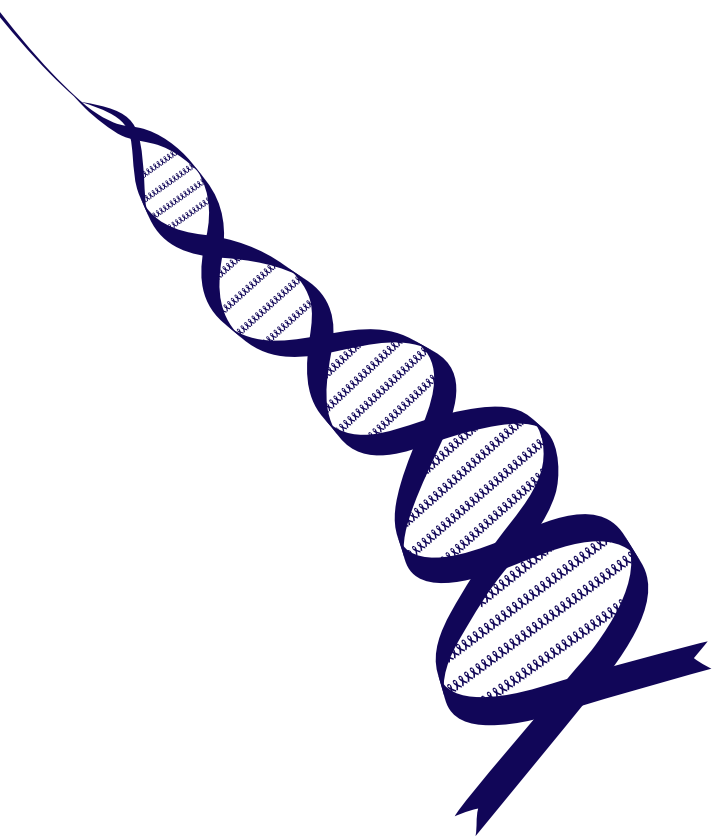
Contents



Chapter 1

General Introduction

E.J. Blok



Breast cancer epidemiology

With an annual incidence of 17.000 per year in the Netherlands, and 1.7 million worldwide, breast cancer is the most frequent malignancy in females, and the second most frequent cancer overall after lung cancer.[1] Over the past years, there is a steady increase in incidence, together with steadily improving survival rates. A major factor in the improving survival rates is the concept of (neo)adjuvant therapy.[2]

Currently, there are three cornerstones for the treatment of primary breast cancer: surgery, radiotherapy and systemic (neo)adjuvant therapy. The aim of surgery is to remove the primary tumour, thereby preventing future metastasis and locoregional complications. The aim of radiotherapy is the prevention of local and regional relapse, especially after breast-conserving surgery and in the presence of lymph node metastasis. The aim of adjuvant systemic therapy, either endocrine therapy, targeted therapy or chemotherapy, is to prevent relapse, in particular distant metastasis. In addition, the concept of neo-adjuvant therapy, in which systemic therapy is applied before the operation, also leads to tumour shrinkage. This could allow breast-conserving surgery, as well as axillary downstaging and provides knowledge about the sensitivity of the tumour towards the therapy.

Approximately 80% of all breast cancers are hormone-receptor positive, meaning that the tumour is expressing either the estrogen receptor (ER), progesterone receptor (PgR) or both.[3] With this expression, the tumour is capable of stimulating its own growth by receiving a higher amount of estrogen-dependent growth signals.[4]

Endocrine therapy

Already in the 19th century, the concept of endocrine sensitivity of breast cancer was discovered, when Col. Sir George Thomas Beatson performed an oophorectomy at three patient with advanced breast cancer, thereby reducing their metastases.[5] For decades, the oophorectomy became standard therapy for advanced breast cancer. In the 1960's and 70's, a first type of chemical endocrine therapy was first discovered. Tamoxifen, a selective estrogen receptor modulator (SERM) first developed as an oral anticonceptive, was shown to have growth inhibiting capacities in HR-positive breast cancer. Initially, this treatment was only used in advanced disease, to inhibit the growth of metastases. However, soon it was also discovered that preventive use of tamoxifen was able to lower the chance of disease recurrence.[6-9] It was established that 2 years, and later 5 years of tamoxifen was associated to a lower rate of recurrences.[10, 11]

In a post-hoc meta-analysis by the Early Breast Cancer Trialists Collaborative Group (EBCTCG), it was established that at 15 years after diagnosis, there was an absolute benefit of 12% on recurrence-free survival of tamoxifen 5 years (33% recurrence) versus no adjuvant endocrine therapy (45% recurrence).[12]

Meanwhile, a second class of endocrine therapy was being developed. Brodie et al first showed the concept of aromatase inhibition, in which the enzyme aromatase, which is responsible for the conversion from androgens to estrogens, is being inhibited.[13] In case of a postmenopausal patient, for which aromatase-dependent androgen conversion is the only source of estrogens, this would lead to a theoretical full depletion of estrogen, thereby preventing any activation of the estrogen receptor.

The Intergroup Exemestane trial (IES trial) was the first trial to directly compare the effect of tamoxifen and aromatase inhibitors, in this case exemestane. After 2-3 years on tamoxifen, patients were randomized between either completing 5 years of tamoxifen, or switching to exemestane to complete 5 years of adjuvant endocrine therapy. Both at 5 and 10 years follow-up, there was a significant improvement for disease-free (DFS) and overall survival (OS) for the switch to exemestane.[14, 15] The Arimidex, Tamoxifen, alone or in combination trial (ATAC) compared anastrozole and tamoxifen monotherapy, together with a third arm combining both agents. The combination arm was closed due to a lack of additional benefit. This study showed that for DFS and distant metastasis-free survival, there was a benefit of an AI over tamoxifen monotherapy.[16, 17] These results were confirmed by the Breast International Group (BIG) 1-98, which additionally showed that a sequential therapy of tamoxifen followed by an AI, was also superior over tamoxifen monotherapy.[18]>

The Tamoxifen and Exemestane Adjuvant Multicenter Trial (TEAM) compared exemestane monotherapy with a sequential scheme of tamoxifen for 2.5 years, followed by exemestane to complete 5 years of therapy. Both at 5 and 10 years of follow-up, this trial showed no difference in DFS or OS.[19, 20] In a recent meta-analysis performed by the EBCTCG, it was confirmed that there was a clinically and statistically significant benefit of both sequential therapy and AI monotherapy over tamoxifen monotherapy, both for DFS and OS. Moreover, a marginal benefit of AI monotherapy over sequential therapy was shown for DFS with an absolute risk reduction of 0.7% (HR 0.90, p=0.045), but not for OS (HR 0.89, p=0.11).[21]

Despite the success of adjuvant endocrine therapy, there is still a continuous risk for recurrences up to 15 years after diagnosis.[22] Therefore, the concept of extended (i.e. longer than 5 years) adjuvant endocrine therapy was developed. Initially, extended therapy was mainly studied in the context of extended 5 years of tamoxifen. After 5 years of tamoxifen, it was shown that extended therapy with either another 5 years of tamoxifen, or 5 years of an AI was beneficial in terms of disease-free survival, in particular in patients with node-positive disease.[23-25] However, the effect on overall survival was marginal, and not statistically significant when data were pooled in a meta-analysis.[26] After receiving an AI in the first 5 years of treatment, until date no study has shown that extended therapy has a significant benefit.

Biomarkers

In all studies mentioned above, it is apparent that there is only a small group of patients that benefits from (extended) adjuvant endocrine therapy. For example, when comparing 5 years of tamoxifen with no adjuvant therapy, the absolute reduction of 12% and the hazard ratio of 0.73 looks impressive, but also means that 88% of the patients is treated in vain.[12] This means that there is a lot of room for improvement. The guidelines for adjuvant therapy are relatively strict; both adjuvant chemotherapy and endocrine therapy are indicated quite easily, thereby lowering the risk for undertreatment. However, due to this approach many patients will be overtreated, which is especially concerning given the side effect profiles of adjuvant chemotherapy and endocrine therapy. Therefore, tailoring strategies are crucial in order to optimize adjuvant therapy, so that maximal benefits can be achieved with minimal harms.

Currently, the possibilities to tailor adjuvant (extended) therapy are limited. For adjuvant chemotherapy, the current tailoring strategies are mainly risk-based. In theory, patients with a higher risk of recurrence will benefit more from adjuvant chemotherapy compared to patients with a low risk of recurrence. Traditionally, clinicopathological factors like lymph node status, tumour size, receptor status, differentiation grade and age are used to determine the risk for recurrence. However, more recently, gene expression profiles have shown to be interesting prognostic tools, accurately predicting the risk for recurrence.[27, 28]

Besides the risk-based tailoring approach, there is also a biology-based approach to tailor therapy. For endocrine therapy, ER and PgR are used as biomarkers to tailor

endocrine therapy. However, since 80% of all breast cancer patients is either ER and/or PgR-positive, and all above-mentioned trials had ER and/or PgR-positivity as inclusion criterion, the level of tailoring reached with only ER and PgR is insufficient.

One factor that could provide a biology-based tailoring approach, next to the hormone receptors, is the immune system. It is well known that the immune system, and the adaptation of the tumour to the immunological burden, is crucial in the development, growth and metastasis of a tumor.[29] One of the many important aspects in the tumour-immune system interaction, are tumour infiltrating lymphocytes (TILs). Although many different subcategories of lymphocytes are present within a tumour, T-cells, and specifically CD8-positive cytotoxic T-cells, are the most abundant. It has been shown for triple-negative breast cancer, and HER2-positive breast cancer, that higher levels of TILs are associated to a better prognosis, and a higher success rate of adjuvant chemotherapy and HER2-targeted therapy.[30-37] Remarkably, this association was not observed for ER-positive breast cancer, in which TILs have no prognostic value.[38, 39] However, the predictive value of TILs on endocrine therapy has never been studied.

In this thesis we aimed to tailor adjuvant therapy, and in particular adjuvant endocrine therapy, using multiple approaches. We particularly focused on the benefits of extended endocrine therapy, and on new translational approaches for tailoring adjuvant therapy for postmenopausal patients with early breast cancer.

Outline of this thesis

Part I of this thesis is aimed at the clinical use of extended endocrine therapy.

Chapter 2 provides an overview of all the current evidence for extended therapy, and the prospectives of the trials that will be reported in the future. Chapter 3 describes the primary results of the multicentre phase III IDEAL trial, in which patients who received any kind of 5 years adjuvant endocrine therapy, were randomized between either 2.5 or 5 years of extended therapy. Chapter 4 provides a more detailed subgroup analysis of the IDEAL trial, trying to identify a clinicopathological subgroup for which there is a benefit of longer extended therapy. Chapter 5 describes which factors are associated to choosing to participate in the IDEAL trial, which factors contributed to treatment compliance, and the impact of treatment compliance on survival.

Part II of this thesis is aimed at tailoring both (neo)adjuvant chemotherapy and endocrine therapy, using biomarker-based approaches.

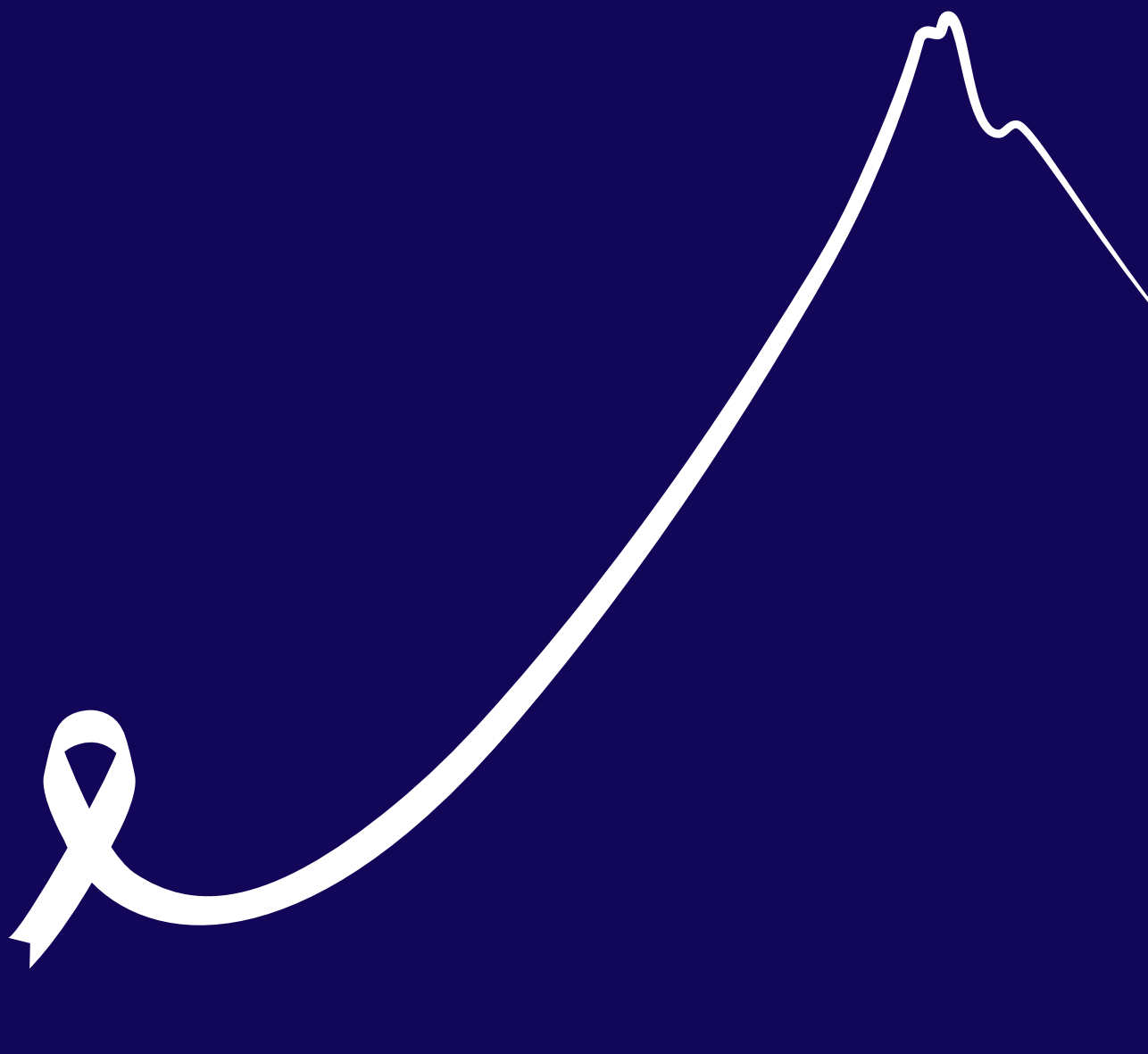
Chapter 6 describes the development and validation of a new platform, capable of assessing the activity of the estrogen receptor pathway. This platform was developed by determining the activity of downstream gene targets of the estrogen receptor, and thereby estimating the probability that the estrogen receptor is active. When the receptor is indeed active, it might predict for the benefit of endocrine therapy, whereas with an inactive pathway endocrine therapy would have little effect. Chapter 7 explores the use of tumour-infiltrating lymphocytes (TILs) as a prognostic marker in ER-positive breast cancer, and as a predictive marker for endocrine therapy. Chapter 8 further studies TILs as a prognostic marker, in combination with expression of the cell surface death receptor FAS. Chapter 9 provides a systematic review of gene expression profiles, in which the four major assays (Endopredict , MammaPrint, OncotypeDX, and Prosigna) are reviewed on the developmental procedure, prognostic and predictive capacities, clinical utility and the economic value of the tests. Chapter 10 provides a critical interpretation of the results of the MINDACT trial, a major clinical trial evaluating the use of MammaPrint in clinical practice.

Finally, a summary and discussion on the results of this thesis will be provided in chapter 11, addressing the future perspectives of tailoring adjuvant therapy.

References

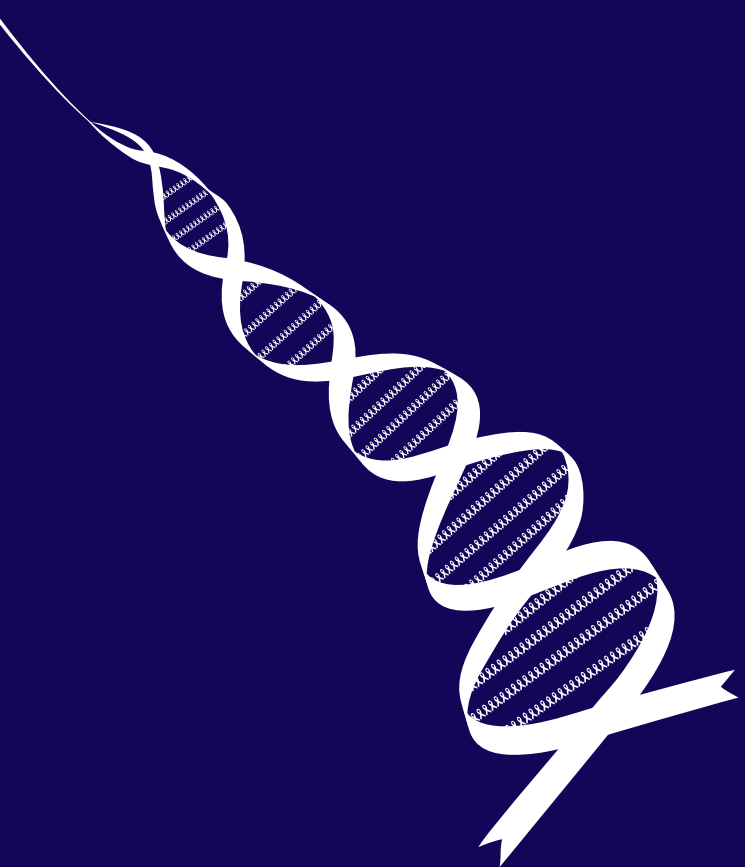
1. Ferlay J, Soerjomataram IF, Dikshit RF et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.(1097-0215 (Electronic)).
2. Narod SA, Iqbal J, Miller AB. Why have breast cancer mortality rates declined? Journal of Cancer Policy 2015;5:8-17.
3. Pritchard KI, Gelmon KA, Rayson D et al. Endocrine therapy for postmenopausal women with hormone-receptor positive her2-negative advanced breast cancer after progression or recurrence on nonsteroidal aromatase inhibitor therapy: a Canadian consensus statement. Current Oncology; Vol 20, No 1 (2013) 2013.
4. Osborne CK, Schiff R. Estrogen-Receptor Biology: Continuing Progress and Therapeutic Implications. JCO 2005;23(8):1616-1622.
5. Beatson G. ON THE TREATMENT OF INOPERABLE CASES OF CARCINOMA OF THE MAMMA: SUGGESTIONS FOR A NEW METHOD OF TREATMENT, WITH ILLUSTRATIVE CASES. The Lancet 1896;148(3803):162-165.
6. CONTROLLED TRIAL OF TAMOXIFEN AS ADJUVANT AGENT IN MANAGEMENT OF EARLY BREAST CANCER. The Lancet 1983;321(8319):257-261.
7. Fisher B, Costantino J, Redmond C et al. A Randomized Clinical Trial Evaluating Tamoxifen in the Treatment of Patients with Node-Negative Breast Cancer Who Have Estrogen-ReceptorGÇôPositive Tumors. N Engl J Med 1989;320(8):479-484.
8. Effects of Adjuvant Tamoxifen and of Cytotoxic Therapy on Mortality in Early Breast Cancer. N Engl J Med 1988;319(26):1681-1692.
9. Fisher B, Costantino J, Redmond C et al. A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor-positive tumors. N Engl J Med 1989;320:479-484.
10. Sacco M, Valentini M, Belfiglio M et al. Randomized Trial of 2 Versus 5 Years of Adjuvant Tamoxifen for Women Aged 50 Years or Older With Early Breast Cancer: Italian Interdisciplinary Group for Cancer Evaluation Study of Adjuvant Treatment in Breast Cancer 01. JCO 2003;21(12):2276-2281.
11. Tamoxifen for early breast cancer: an overview of the randomised trials. The Lancet 1998;351(9114):1451-1467.
12. Early Breast Cancer Trialists' Collaborative Group (. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-1717.
13. Jordan VC, Brodie AMH. Development and Evolution of Therapies Targeted to the Estrogen Receptor for the Treatment and Prevention of Breast Cancer. Steroids 2007;72(1):7-25.
14. Coombes RC, Hall E, Gibson LJ et al. A Randomized Trial of Exemestane after Two to Three Years of Tamoxifen Therapy in Postmenopausal Women with Primary Breast Cancer. N Engl J Med 2004;350(11):1081-1092.
15. Morden JP, Alvarez I, Bertelli G et al. Long-Term Follow-Up of the Intergroup Exemestane Study. JCO 2017;JCO.
16. Baum M, Budzar AU, Cuzick J et al. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. Lancet 2002;359:2131-2139.
17. Cuzick J, Sestak I, Baum M et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. Lancet Oncol 2010;11:1135-1141.
18. The BIG 1-98 Collaborative Group. Letrozole Therapy Alone or in Sequence with Tamoxifen in Women with Breast Cancer. N Engl J Med 2009;361(8):766-776.
19. Derks MGM, Blok EJ, Seynaeve C et al. Adjuvant tamoxifen and exemestane in women with postmenopausal early breast cancer (TEAM): 10-year follow-up of a multicentre, open-label, randomised, phase 3 trial. The Lancet Oncology 2017;18(9):1211-1220.
20. van de Velde CJ, Rea D, Seynaeve C et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-331.
21. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.

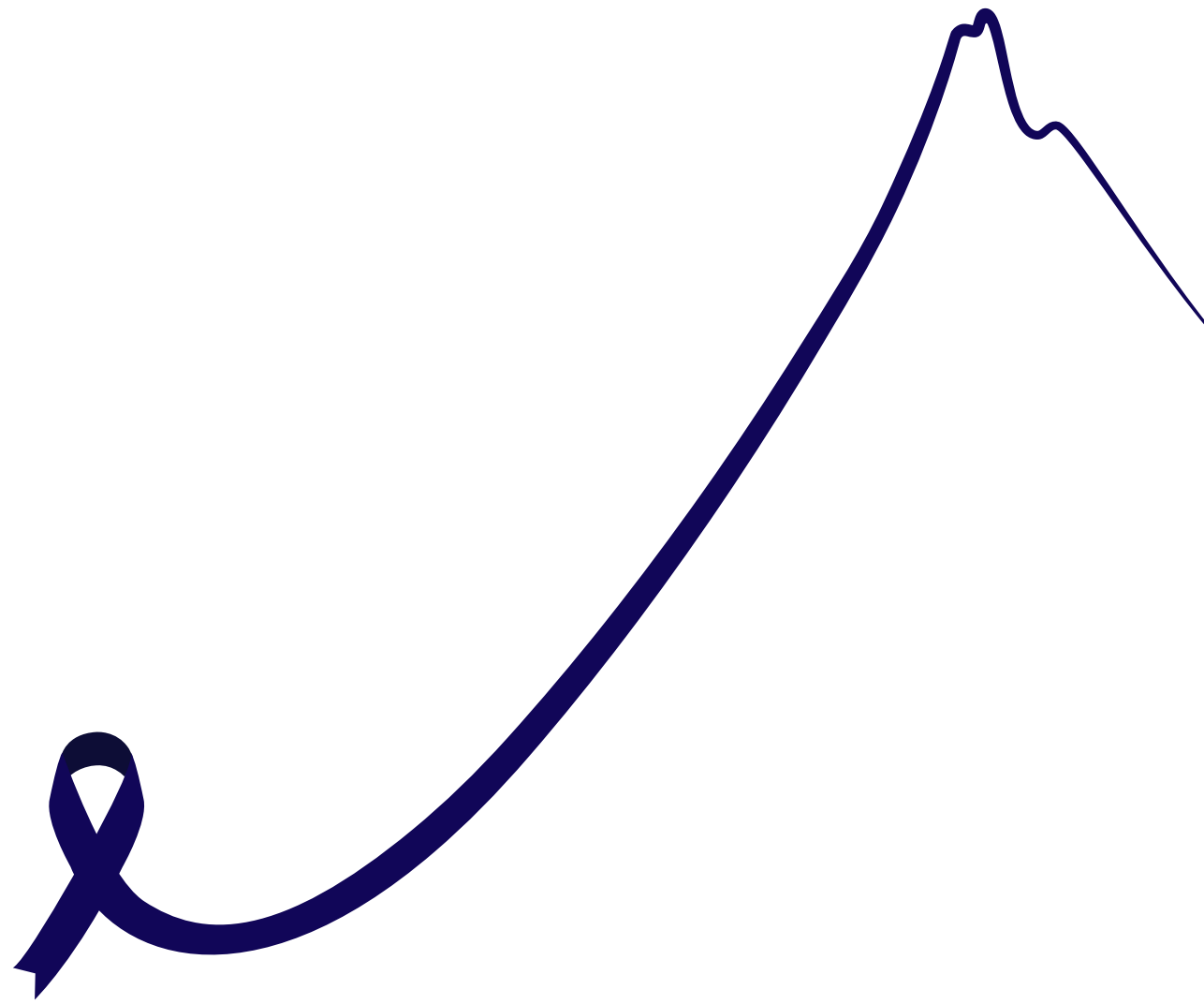
22. Pan H, Gray RG, Davies C et al. Predictors of recurrence during years 5-14 in 46,138 women with ER+ breast cancer allocated 5 years only of endocrine therapy (ET). JCO 2016;34(15_suppl):505.
23. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381:805-816.
24. Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).
25. Goss PE, Ingle JN, Martino S et al. A Randomized Trial of Letrozole in Postmenopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. N Engl J Med 2003;349(19):1793-1802.
26. Ibrahim EM, Al-Hajeili MR, Bayer AM et al. Extended adjuvant endocrine therapy in early breast cancer: a meta-analysis of published randomized trials. Medical Oncology 2017;34(7):131.
27. Cardoso F, vant Veer LJ, Bogaerts J et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med 2016;375(8):717-729.
28. Sparano JA, Gray RJ, Makower DF et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. N Engl J Med 2015;373(21):2005-2014.
29. de Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. Nat Rev Cancer 2006;6(1):24-37.
30. West NR, Kost SE, Martin SD et al. Tumour-infiltrating FOXP3(+) lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. Br J Cancer 2013;108(1):155-162.
31. West NR, Milne K, Truong PT et al. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res 2011;13(6):R126.
32. Denkert C, Loibl S, Noske A et al. Tumor-Associated Lymphocytes As an Independent Predictor of Response to Neoadjuvant Chemotherapy in Breast Cancer. JCO 2010;28(1):105-113.
33. Loi S, Sirtaine N, Piette F et al. Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. JCO 2013;31(7):860-867.
34. Loi S, Michiels S, Salgado R et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Annals of Oncology 2014;25(8):1544-1550.
35. Salgado R, Denkert C, Demaria S et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. Annals of Oncology 2015;26(2):259-271.
36. Ali HR, Provenzano E, Dawson SJ et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. Annals of Oncology 2014;25(8):1536-1543.
37. Adams S, Gray RJ, Demaria S et al. Prognostic Value of Tumor-Infiltrating Lymphocytes in Triple-Negative Breast Cancers From Two Phase III Randomized Adjuvant Breast Cancer Trials: ECOG 2197 and ECOG 1199. JCO 2014.
38. Mahmoud SMA, Paish EC, Powe DG et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. JCO 2011;29(15):1949-1955.
39. Baker K, Lachapelle J, Zlobec I et al. Prognostic significance of CD8+ T lymphocytes in breast cancer depends upon both oestrogen receptor status and histological grade. Histopathology 2011;58(7):1107-1116.



Part 1

Clinical use of extended endocrine therapy



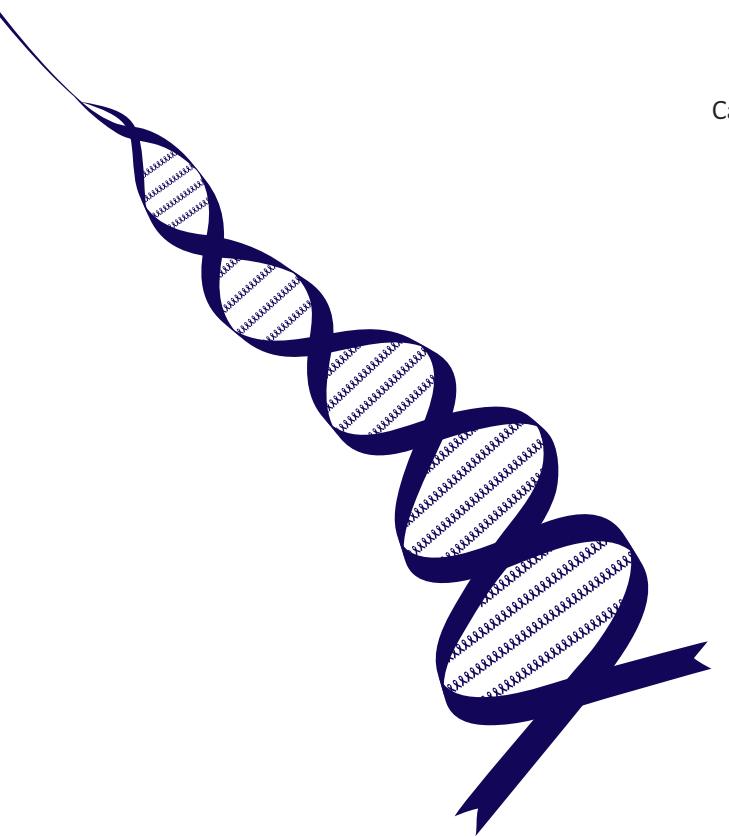


Chapter 2

Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: current and future evidence

E.J. Blok
M.G.M. Derks
J.J.M. van der Hoeven
C.J.H. van de Velde
J.R. Kroep

Cancer Treat Rev 2015; 41(3): 271-276



Abstract

The optimal duration and regimen of adjuvant hormonal therapy for premenopausal and postmenopausal patients with hormone receptor positive early breast cancer has not yet been established. This review will give an overview of published and ongoing studies concerning extended endocrine treatment. Most of the currently published studies are based on the adjuvant treatment regime of 5 years tamoxifen, which has been proven to be inferior compared to aromatase inhibitor (AI)-containing regimes. Therefore, until today, there is no clear evidence for the extension of endocrine therapy after upfront AI-based adjuvant treatment regimes. Multiple clinical trials, which will be discussed in this review, are ongoing to elucidate on this matter. We emphasize the need for tailoring of extended adjuvant endocrine treatment. The quest for predictive biomarkers, which are currently being investigated in the context of decision-making whether or not to start adjuvant chemotherapy, should be expanded to include the feasibility of extended endocrine treatment based on these markers. By tailoring the extension of endocrine treatment, overtreatment, side effects and unnecessary costs will be prevented.

Introduction

Nowadays, endocrine treatment is one of the mainstays of breast cancer treatment, but the optimal duration is yet to be determined. It is estimated that 75% of all breast cancer patients are hormone receptor-positive (HR+) breast cancer, and therefore might benefit from endocrine treatment.[1] Endocrine therapy significantly reduced the risk of death among patients with HR-positive tumors compared to those with ER and PgR-negative tumors. Five years of adjuvant tamoxifen reduced the breast cancer mortality by about a third throughout the first 15 years.[2] However, estimations for the long term risk of recurrence show that HR- positive breast cancer patients remain at a significant risk of recurrence until at least 15 years post diagnosis, whereas the risk for recurrence for ER/PgR negative patients is highest shortly after diagnosis but decreases below that of ER/PgR positive patients later on.[2,3] There is scientific evidence that it is beneficial to use extended tamoxifen after 5 years of adjuvant tamoxifen[4,5] and to start using an aromatase inhibitor after having received tamoxifen for 5 years, even if tamoxifen was stopped a considerable time ago.[6]

Adjuvant endocrine therapy

Ever since the first oophorectomy performed by dr. Beatson in 1896[7], endocrine therapy has been established as a treatment option for HR+ breast cancer. Currently, tamoxifen and aromatase inhibitors (AIs) are the two most important categories for endocrine treatment in postmenopausal patients. A third category of endocrine therapy, ovarian function suppression (OFS by GnRH agonists, ablation or radiotherapy) is used in premenopausal patients to diminish the ovarian function in combination with tamoxifen or AIs.[8]

After its introduction in 1970, the selective estrogen receptor antagonist tamoxifen soon became standard therapy in the treatment of advanced hormone receptor-positive breast cancer9. Initially, treatment was based on 1-2 year strategies as this was the optimal duration in advanced disease.[9,10] However, it became clear that 5 year adjuvant treatment improved the clinical outcome, and for decades this has been the standard treatment for hormone receptor-positive breast cancer.[11-13] Five years of adjuvant treatment with tamoxifen versus no treatment showed a relative risk reduction in 15 year recurrence risk of 40%, with an absolute gain of 13.2%.[2] Furthermore, a decrease of 15 year breast cancer mortality has been observed with a relative risk of 0.7, and an absolute benefit of 9.2%.

While tamoxifen was introduced, the first AIs were developed and proven to be efficient in metastatic breast cancer patients.[14] However, due to its inhibitory function on cytochrome P450, its effects on adrenal function and subsequent side effects, the first and second generation AIs did not become mainstream treatment for adjuvant treatment, and were only used in separate cases of metastatic disease.[14,15] AIs only became popular after the development of third generation compounds (anastrozole, letrozole and exemestane) which are less toxic. The first report of these third generation AIs in the setting of a large clinical trial was in the Anastrozole, Tamoxifen Alone or in Combination (ATAC) trial, in which anastrozole, tamoxifen and a combination of both were studied16. At initial 5 years and 10 years follow-up, this study showed the superiority of AIs over tamoxifen as a first line adjuvant treatment for early breast cancer in postmenopausal patients, and comparable results for the combination treatment.[16-18] After these findings, multiple trials examined the effect of switching to an AI compared to continuing with tamoxifen. A meta-analysis by Dowsett *et al* in 2010 showed a superiority of this switch scheme above continuing with tamoxifen.[19] This switch scheme consists of 2-3 years of tamoxifen, followed by 2-3 years of an AI. Two other major trials, BIG 1-98 and TEAM-trial, initially focused on the same research question whether AIs would be superior to tamoxifen. However, due to the results that AIs appeared superior to tamoxifen, they changed their design into a comparison of five years AI with the before-mentioned switch scheme. Both studies showed a borderline non-significant progressive decrease of disease free survival (DFS) or recurrence free survival (RFS) in the initial 2-3 years of tamoxifen. However, after the switch to an AI, the difference between both groups stabilized leading to a non-significant difference between both groups.[20,21] Therefore there is no evidence for superiority of either 5 years AIs or a switch scheme at long term follow-up.

For premenopausal patients monotherapy with tamoxifen was the standard therapy for a long time with a possible benefit from ovarian suppression for patients of 40 year and younger.[22,23] Recently, the results of the TEXT and SOFT trial revealed that for premenopausal patients addition of ovarian function suppression should be considered for patients younger than 35 years (5 year breast cancer free interval of 67.7% for tamoxifen vs 78.9% for tamoxifen plus OFS and 83.4% for exemestane plus OFS) or who received chemotherapy (5 year breast cancer free interval 78% for tamoxifen vs 82.5% for tamoxifen plus OFS vs 85.7% for exemestane plus OFS.[24]

Side effects of aromatase inhibitors are different from those of tamoxifen. Generally, tamoxifen is well tolerated, with most reported events to be hot flushes, osteoporosis, arthralgia and gynaecologic symptoms like vaginal bleeding and discharge.[17] More severe toxicities which have been described with the use of tamoxifen are venous thromboembolisms and a hazard ratio of approximately 2 for endometrial carcinomas and mood change or depression.[25-29] For aromatase inhibitors, hypertension, dyslipidaemia, arthralgia and osteoporosis are more frequently described. Gynaecological symptoms and hot flushes are less common.[16,17,30-32] Arthralgia is usually reported by patients as the most relevant side effect.[30,33] Generally, just as tamoxifen, aromatase inhibitors are relatively well tolerated. In designated trials comparing the switch scheme with aromatase inhibitors only, no important differences in side effects or quality of life were shown.[34] The TEAM trial showed that in general, there are more gynaecological and vascular side effects with the tamoxifen-containing switch scheme, while in the aromatase inhibitor group hypertension, dyslipidaemia and musculoskeletal complaints were more pronounced.[20] Similar results were observed in the BIG 1-98 study.[21] Therefore, regarding side effects and toxicity, therapy choices should be tailored on the individual patient taking co-morbidity and patients preference in consideration.

These findings have led to the conclusion that AIs should be included in the adjuvant treatment of early HR+ breast cancer in postmenopausal patients, and also in combination with ovarian suppression for premenopausal patients. However, there is no evidence for superiority of either 5 years aromatase inhibitors or a switch scheme of tamoxifen followed by an AI. This review will comment on the current evidence for therapy extension, ongoing studies and possible predictive markers suitable for decision-making concerning extended endocrine treatment.

Extended therapy

The current period of 5 or 10 years of adjuvant endocrine treatment for early breast cancer is based on early results of adjuvant tamoxifen.[2,5,35] However, it was shown that approximately 50% of recurrences happened after the initial 5 years of adjuvant treatment.[2,36] These findings initiated a debate on the optimal duration of therapy, and a number of studies was set up to elucidate on this matter.

The NCIC CTG MA.17 trial in 5187 patients showed that 10 years of treatment (5 years of tamoxifen followed by 5 years of letrozole) was superior to five years of tamoxifen.[6]

Table 1 – evidence for extended therapy

study	Nr of patients	initial therapy	extended treatment	control arm	FU time	HR DFS	p-value	HR OS	p-value	reference
MA.17	5187	5y TAM	5y letrozole	placebo	30	0.57	<0.0001	0.76	0.25	6
ABCSG	860	5y TAM	3y anastrozole	no treatment	62.3	0.62	0.031	0.89	0.57	39
NSABP B33	1598	5y TAM	5y exemestane	placebo	30	0.68	0.07	NA	NA	41
ATENA	448	5y TAM	5y exemestane	placebo	NA	NA	NA	NA	NA	40
ATLAS	6846	5y TAM	5y TAM	no treatment	NA	0,75*	0.002	0,87*	0.01	4
aTTom	6953	5y TAM	5y TAM	no treatment	NA	0,75*	NA**	0,86*	NA**	5

All major clinical trials on extended adjuvant endocrine treatment that have been published are shown in this table. For some studies, no data were available in the original publication (NA = not available).

* at >10y follow up

** no p-values provided, but significant based on confidence interval

After a median follow-up of 30 months, a hazard ratio of disease free survival of 0.58 was found, with a non-significant HR of 0.76 for overall survival. Upon these interim results the study was unblinded, and cross-over was allowed. However, with a 66% cross-over from the placebo to treatment arm, there was a significant loss of power for further follow-up. At 60 months of follow up, this has led to a HR for disease free survival of 0.68 (0.55-0.83), but no difference in overall survival (HR=0.98). With a statistical test called the inverse probability of censoring weighted analysis (IPCW-analysis), they estimated that the HR for overall survival would have been 0.61 (0.52-0.71) without cross-over.[6,37,38] Although this was the first proof of principle for extended endocrine therapy, the interpretation of these findings is difficult. Starting five years of letrozole after 5 years of tamoxifen is basically the same strategy as the switch scheme described above, only with longer treatment intervals. It could be stated that this study confirms the benefits of a (late) switch scheme, rather than a general benefit for extended therapy. In 2006, Ingle *et al* showed that the hazard ratios for disease free survival when using letrozole decreased over time, which was attributed to an increasing risk of recurrence in the placebo-controlled group.[39] These findings indicate a possible benefit for extending the treatment even further beyond the studied term of 5 years. Whether this also implies for patients who received up-front AI treatment is only supported by circumstantial evidence, and has not been studied yet.

Three other, smaller studies have confirmed the results of the MA.17 study (Table 1). The Austrian Breast and Colorectal Cancer Study Group (ABCSG)-6a study, had a similar setup in which 856 patients after 5 years of tamoxifen were randomized between 3 years of anastrozole or regular follow-up.[40] A reduction of 38% in the risk of breast cancer recurrence was observed (HR 0.62, 95% CI 0.4-0.96), which is in concordance with the MA.17 results. This study failed to show any benefit on overall survival, most likely due to the relative short follow-up of 5 years. Two other studies, both evaluating exemestane as extended therapy after 5 years of tamoxifen, were closed prematurely due to the results of the MA.17 trial.[41,42] One of them however published their underpowered results, already showing a borderline significant decrease in DFS at 30 months of follow-up42. A meta-analysis conducted with the four trials mentioned above, has led to an overall decrease in breast cancer recurrence of 43% (absolute decrease of 2.9%) and a (not statistically significant) decrease in mortality of 11% (absolute decrease of 0.5%) at 2.5 years of follow-up.[43] The consistent results in these four trials using letrozole, anastrozole and exemestane as AIs, lead to

another conclusion that the advantage of AIs is not limited to one specific type, but is a class effect. There appears to be no difference between the separate agents, making future comparisons and meta-analyses less complicated.

Extended therapy after 5 years of tamoxifen has comprehensively been studied. Early small studies did not show a benefit for extended tamoxifen, with an increase in toxicity. In 2013, the 15-year follow-up results from Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) trial were published4. This study, which randomized nearly 7000 ER-positive patients between 5 or 10 years tamoxifen, showed a benefit for continuing tamoxifen with an absolute benefit of 3.7% (21.4% vs 25.1%) on recurrence risk, and an absolute mortality reduction of 2.8% (12.2% vs 15%). Remarkably, these benefits were mainly observed in the period after 10 years when treatment was ceased. This was attributed to a carryover effect, which is well known for tamoxifen.[2] Similar results were observed in the British Adjuvant Tamoxifen - To Offer More (ATTOM) trial.[5]

Clinical implications

The interpretation of these extended therapy studies is difficult. As shown in Table 1, all available studies are based on 5 years of tamoxifen, before therapy was extended. Extended tamoxifen has shown a small but consistent overall survival benefit. The results of the ATAC, BIG 1-98 and TEAM trial clearly show that AI-containing adjuvant regimes, either as a monotherapy or as a switch-scheme, are preferred above tamoxifen monotherapy. As a result, there is no clear evidence for therapy extension of modern 'regular' AI-containing adjuvant treatment, and no direct evidence for extended therapy with an AI longer than 5 years. Furthermore, there are no studies available with a direct comparison between the extension with tamoxifen or letrozole.

The most recent ASCO guidelines, published in July 2014, support –based on recent literature data-multiple treatment strategies for the type and length of adjuvant endocrine therapy.[44] They offer four options: tamoxifen for 10 years, tamoxifen for 5 years followed by an AI for 5 years, AI for 5 years or a switch scheme starting with tamoxifen for 2-3 years followed by an AI for up to 5 years. Little evidence is available to compare these four options. Only a comparison between the switch scheme and 5 years of AI is available, which has led to no significant differences as discussed earlier. In a review published in 2013, Strasser-Weippl *et al* performed an unofficial analysis comparing extended therapy using AIs with tamoxifen after 5 years of tamoxifen. Comparing hazard ratios of two separate studies, they state that switching to an AI

after 5 years of tamoxifen appears beneficial over continuing with tamoxifen, and that it would lead to a larger recurrence rate reduction and a better overall survival.[45] Although comparing hazard ratios of different studies is controversial, this would be in accordance with the findings in 'regular' adjuvant treatment that AI-containing regimes have better outcomes compared to tamoxifen monotherapy.

For premenopausal women the evidence based choices are: tamoxifen 5-10 years, tamoxifen 5 years followed by AI 5 years, ovarian suppression with tamoxifen or AI which should be considered for higher risk patients (<35 years, premenopausal after prior chemotherapy and multiple positive axillary nodes). The optimal duration of ovarian suppression based therapy is uncertain; the SOFT and TEXT trial both studied 5 years.

Ongoing studies

Now that the first studies have reported a benefit of extended adjuvant endocrine therapy in early breast cancer, many challenges lie ahead. Basically, it comprises three main topics: (1) To validate the findings of earlier studies in modern AI-containing adjuvant therapy, (2) to determine the optimal duration of extended therapy and (3) to identify and further explore predictive factors for patients that would benefit most of extended therapy46. Figure 1 summarizes the ongoing and unpublished trials.[37,47,48]

The main study focussing on validation of extended therapy after modern, AI-containing, regimes is the NSABP B42 trial, in which nearly 4000 patients were randomized between 5 years of letrozole or placebo, after 5 years of regular adjuvant therapy either consisting of aromatase inhibitors or tamoxifen followed by aromatase inhibitors.[49] Also the Letrozole Adjuvant Therapy Duration (LEAD) and the Different Durations of Anastrozole after Tamoxifen (DATA) trials have the same perspective, by randomizing patients after 2-3 years of adjuvant tamoxifen between standard treatment (additional 2-3 years AI) and extended treatment (5-6 years of AI), respectively.[50,51] This creates a situation in which standard therapy is compared with 2.5 years extended therapy. Both the Secondary Adjuvant Long-term Study with Arimidex (SALSA; ABCSG-16) and Investigation on the Duration of Extended Adjuvant Letrozole treatment (IDEAL) randomize between 2-2.5 and 5 years of therapy extension, after any prior adjuvant endocrine treatment of 5 years.[52,53] The Study of Letrozole Extension (SOLE) trial compares 5 years of continuous AI therapy extension with intermittent letrozole extension. This intermittent scheme consists of

an annual cycle of 9 months therapy and a 3 months break, while the final (5th) year is 12 months of therapy.[54] Combined, these studies will presumably answer whether therapy extension after 5 years of AI-containing adjuvant treatment is relevant, and what would be the optimal duration of extension (2.5 vs 5 years). Finally, there is also an extension of the MA.17 trial, called the MA17R.[37] This study will extend the adjuvant therapy even further with another 5 years of letrozole versus placebo, which counts up towards 15 years of adjuvant therapy.

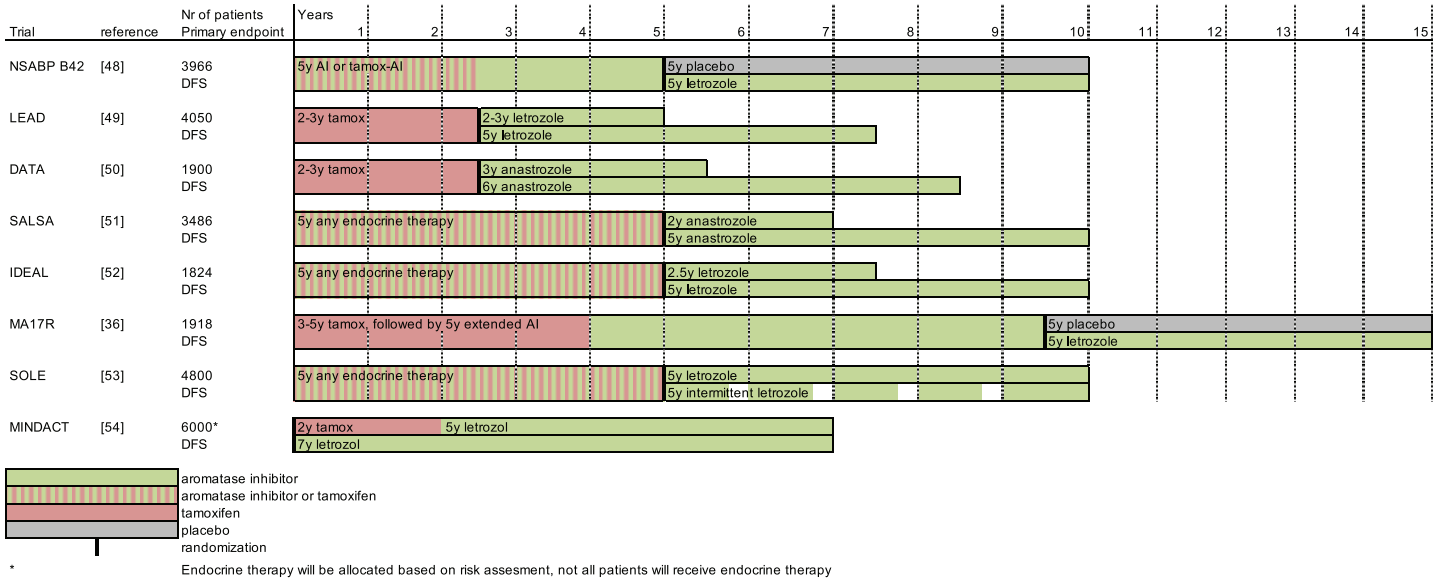
The MINDACT-trial, a large trial with the main purpose on decision making based on either epidemiological data or genetic data, also has a substudy in which 7 years of AI is compared to 2 years of tamoxifen followed by 5 years of AI.[55] Although this study is not focused on the optimal duration of therapy (both arms have the same treatment length), it will add more value to extended AI beyond 5 years.

Predictive markers

Because luminal breast cancer is a heterogeneous disease it is important to be able to select those patients that will benefit most from extended adjuvant therapy. For adjuvant endocrine treatment, many efforts have been made to identify biomarkers or molecular profiles capable of predicting response to endocrine treatment and the risk of recurrence after treatment.[56,57] Although research is still ongoing, established methods comprise both immunohistochemical and genetic approaches. However, most of these platforms are only validated for use in regular endocrine therapy, or for adjuvant chemotherapy. None of them is validated for extended endocrine treatment.

Classical risk factors like age and nodal status were analysed in the MA.17 trial, both showing no statistical differences between the subgroups.[6,37] Regular immunohistochemical (IHC) markers in breast cancer comprise the Estrogen Receptor (ER), Progesterone Receptor (PgR) and the Human Epidermal growth factor Receptor 2 (HER2). For the ER receptor it was shown in the TEAM trial, that a semi-quantitative expression analysis using IHC is predictive for adjuvant endocrine therapy response.[58] HER2, initially discovered as a predictive marker for poor prognosis and later on developed as a target for monoclonal antibodies against the Her2 receptor, such as trastuzumab (Herceptin®), was also associated with resistance against endocrine treatment.[59-61] Ki-67 also showed to be predictive for the response on endocrine treatment, and the difference in Ki-67 measurement after 2 weeks of neo-adjuvant endocrine therapy appeared to be predictive for the long term effect of endocrine treatment.[62,63]

figure 1 – Overview of ongoing trials



This figure shows the currently ongoing trials investigating extended endocrine treatment. In studies where time ranges instead of time points are used, the expected average treatment duration is plotted. Abbreviations: nr = number, AI = aromatase inhibitor, tam = tamoxifen, y = year, DFS = Disease Free Survival

Based on the immunohistochemical markers mentioned above, a number of multi-marker assays has been developed to predict recurrence risk. IHC4, which consists of a single score based on the expression of ER, PgR, HER2 and ki-67, has been developed as a platform to predict recurrence risk in early breast cancer.[64,65] This assay has been validated retrospectively in the setting of adjuvant endocrine treatment in the ATAC study. IHC4 is of value in clinical decision making, especially in combination with clinicopathologic parameters like tumour grade, size, nodal status and age.[66,67]

A similar platform, called the preoperative endocrine prognostic index (PEPI), was developed specifically for neo-adjuvant treatment. It comprises a combination of post-treatment ER expression, Ki67, histological grade, tumour size, nodal status, and treatment response. This platform was able to stratify patients in three risk groups, with relapse risks of 10%, 23%, and 48%68. The authors suggest that this assay would assist in the decision of starting adjuvant chemotherapy, but it could also be worthwhile to validate this platform for use in decisions concerning endocrine treatment extension.

Both the Mammaprint and Oncotype DX have been established as commercially available genetic testing platforms, depending on the expression of respectively 70 and 21 genes known to be correlated with recurrent disease. Both tests are capable of stratifying the risk of recurrence in low, medium (only for the Oncotype DX) and high. This stratification indirectly represents the likelihood of benefit from chemotherapy. Both tests are currently being validated in a prospective study, with regard to decision-making for adjuvant chemotherapy.[55,69] Another genetic platform called the Endopredict, which was developed as specific for endocrine therapy using eight genes involved in ER-signalling, was used to calculate risk of recurrence after 5 years of endocrine treatment.[70] Furthermore, the Breast Cancer Index (BCI) and Prosigma Risk of Recurrence (ROR), are multigene assays capable of predicting recurrence, although only the ROR-score provided significant prognostic information for late recurrences (5-10 years).[71,72]

To the authors' knowledge, none of these markers and platforms has been validated in a cohort of patients on extended endocrine treatment. It would not be unlikely that all these markers and platforms described, might also be valuable to assist in deciding whether or not to extend endocrine treatment. For this, validation in the setting of extended endocrine treatment is necessary. Although this is an expensive

and elaborative procedure, it is a crucial step towards tailoring of adjuvant endocrine treatment.

Furthermore, research into new predictive and prognostic markers like (epi)genetics, proteomics, circulating tumour DNA, circulating tumour cells (CTCs) and other promising techniques could be valuable in the setting of extension of endocrine treatment.

Conclusion

After almost 5 decades of endocrine therapy, there is still debate on the optimal combination and length of adjuvant therapy. Although studies currently available give strong suggestions that extension of endocrine therapy has benefits, there is actually no strong evidence to support this in the current clinical setting using AIs in the initial adjuvant treatment. Nonetheless, extended endocrine therapy is a promising strategy to further reduce the risk of recurrence. In future studies, emphasis should be laid on selection of subpopulations who benefit most from therapy extension. Patient tailored decision-making will eventually prevent overtreatment, side effects and costs, and add great value to the treatment of breast cancer.

Conflict of interest

All authors have declared that they have no conflicts of interest in regards to this manuscript.

Acknowledgements

This review is an investigator driven study, performed by the department of Surgery and the department of Medical Oncology in the Leiden University Medical Center, using an educational grant provided by Novartis.

References

1 Pritchard KI, Gelmon KA, Rayson D, et al. Endocrine therapy for postmenopausal women with hormone receptor-positive her2-negative advanced breast cancer after progression or recurrence on nonsteroidal aromatase inhibitor therapy: a Canadian consensus statement. Current Oncology; Vol 20, No 1 (2013).

2 Davies C, Godwin J, Gray R, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011 Aug 27;378(9793):771-84.

3 Saphner T, Tormey DC, Gray R. Annual hazard rates of recurrence for breast cancer after primary therapy. Journal of Clinical Oncology 1996 Oct 1;14(10):2738-46.

4 Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381:805-16.

5 Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).

6 Goss PE, Ingle JN, Martino S, et al. A Randomized Trial of Letrozole in Postmenopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. N Engl J Med 2003 Nov 6;349(19):1793-802.

7 Beatson G. ON THE TREATMENT OF INOPERABLE CASES OF CARCINOMA OF THE MAMMA: SUGGESTIONS FOR A NEW METHOD OF TREATMENT, WITH ILLUSTRATIVE CASES. The Lancet 1896 Jul 18;148(3803):162-5.

8 Pagani O, Regan MM, Walley BA, et al. Adjuvant Exemestane with Ovarian Suppression in Premenopausal Breast Cancer. N Engl J Med 2014 Jun 1;371(2):107-18.

9 Cole MP, Jones CT, Todd ID. A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474. Br J Cancer 1971 Jun;25(2):270-5.

10 Early Breast Cancer Trialists' Collaborative Group. Effects of Adjuvant Tamoxifen and of Cytotoxic Therapy on Mortality in Early Breast Cancer. N Engl J Med 1988 Dec 29;319(26):1681-92.

11 Current Trials Working Party of the Cancer Research Campaign Breast Cancer Trials Group. Preliminary Results From the Cancer Research Campaign Trial Evaluating Tamoxifen Duration in Women Aged Fifty Years or Older With Breast Cancer. Journal of the National Cancer Institute 1996 Dec 18;88(24):1834-9.

12 Sacco M, Valentini M, Belfiglio M, et al. Randomized Trial of 2 Versus 5 Years of Adjuvant Tamoxifen for Women Aged 50 Years or Older With Early Breast Cancer: Italian Interdisciplinary Group for Cancer Evaluation Study of Adjuvant Treatment in Breast Cancer 01. Journal of Clinical Oncology 2003 Jun 15;21(12):2276-81.

13 Swedish Breast Cancer Cooperative Group. Randomized Trial of Two Versus Five Years of Adjuvant Tamoxifen for Postmenopausal Early Stage Breast Cancer. Journal of the National Cancer Institute 1996 Nov 6;88(21):1543-9.

14 Smith IE, Fitzharris BM, Mckinna JA, et al. AMINOGLUTETHIMIDE IN TREATMENT OF METASTATIC BREAST CARCINOMA. The Lancet 1978 Sep 23;312(8091):646-9.

15 Dowsett M. Clinical development of aromatase inhibitors for the treatment of breast and prostate cancer. The Journal of Steroid Biochemistry and Molecular Biology 1990 Dec 20;37(6):1037-41.

16 Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. The Lancet 2002 Jun 22;359(9324):2131-9.

17 Howell A, Cuzick J, Baum M, et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. The Lancet 2005 Jan 1;365(9453):60-2.

18 Cuzick J, Sestak I, Baum M, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. The Lancet Oncology 2010 Dec;11(12):1135-41.

19 Dowsett M, Cuzick J, Ingle J, et al. Meta-Analysis of Breast Cancer Outcomes in Adjuvant Trials of Aromatase Inhibitors Versus Tamoxifen. Journal of Clinical Oncology 2010 Jan 20;28(3):509-18.

20 van de Velde CJ, Rea D, Seynaeve C, et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-31.

21 The BIG 1-98 Collaborative Group. Letrozole Therapy Alone or in Sequence with Tamoxifen in Women with Breast Cancer. N Engl J Med 2009 Aug 20;361(8):766-76.

22 Davidson NE, O'Neill AM, Vukov AM, et al. Chemoendocrine Therapy for Premenopausal Women With Axillary Lymph Node-Positive, Steroid Hormone Receptor-Positive Breast Cancer: Results From INT 0101 (E5188). Journal of Clinical Oncology 2005 Sep 1;23(25):5973-82.

23 LHRH-agonists in Early Breast Cancer Overview group. Use of luteinising-hormone-releasing hormone agonists as adjuvant treatment in premenopausal patients with hormone-receptor-positive breast cancer: a meta-analysis of individual patient data from randomised adjuvant trials. The Lancet 2007 May 19;369(9574):1711-23.

24 Pagani O, Regan MM, Walley BA, et al. Adjuvant Exemestane with Ovarian Suppression in Premenopausal Breast Cancer. N Engl J Med 2014 Jun 1;371(2):107-18.

25 Malfetano JH. Tamoxifen-associated endometrial carcinoma in postmenopausal breast cancer patients. Gynecologic Oncology 1990 Oct;39(1):82-4.

26 Magriples U, Naftolin F, Schwartz PE, et al. High-grade endometrial carcinoma in tamoxifen-treated breast cancer patients. Journal of Clinical Oncology 1993 Mar 1;11(3):485-90.

27 Fornander T, Cedermark Br, Mattsson A, et al. ADJUVANT TAMOXIFEN IN EARLY BREAST CANCER: OCCURRENCE OF NEW PRIMARY CANCERS. The Lancet 1989 Jan 21;333(8630):117-20.

28 Duggan C, Marriott K, Edwards R, et al. Inherited and Acquired Risk Factors for Venous Thromboembolic Disease Among Women Taking Tamoxifen to Prevent Breast Cancer. Journal of Clinical Oncology 2003 Oct 1;21(19):3588-93.

29 Meier CR, Jick H. Tamoxifen and risk of idiopathic venous thromboembolism. British Journal of Clinical Pharmacology 1998 Jun 1;45(6):608-12.

30 Garreau JR, DeLaMelena T, Walts D, et al. Side effects of aromatase inhibitors versus tamoxifen: the patients perspective. The American Journal of Surgery 2006 Oct;192(4):496-8.

31 Winer EP, Hudis C, Burstein HJ, et al. American Society of Clinical Oncology Technology Assessment on the Use of Aromatase Inhibitors As Adjuvant Therapy for Postmenopausal Women With Hormone Receptor-Positive Breast Cancer: Status Report 2004. Journal of Clinical Oncology 2005 Jan 20;23(3):619-29.

32 Smith IE, Dowsett M. Aromatase Inhibitors in Breast Cancer. N Engl J Med 2003 Jun 12;348(24):2431-42.

33 Donnellan PP, Douglas SL, Cameron DA, et al. Aromatase Inhibitors and Arthralgia. Journal of Clinical Oncology 2001 May 15;19(10):2767.

34 Cella D, Fallowfield L, Barker P, et al. Quality of Life of Postmenopausal Women in the ATAC (Arimidex, Tamoxifen, Alone or in Combination) Trial after Completion of 5-years' Adjuvant Treatment for Early Breast Cancer. Breast Cancer Res Treat 2006;100(3):273-84.

35 Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. The Lancet 2009;381(9869):805-16.

36 Brewster AM, Hortobagyi GN, Broglio KR, et al. Residual Risk of Breast Cancer Recurrence 5 Years After Adjuvant Therapy. Journal of the National Cancer Institute 2008 Aug 20;100(16):1179-83.

37 Higgins MJ, Liedke PER, Goss PE. Extended adjuvant endocrine therapy in hormone dependent breast cancer: The paradigm of the NCIC-CTG MA.17/BIG 1-97 trial. Critical Reviews in Oncology/Hematology 2013 Apr;86(1):23-32.

38 Jin H, Tu D, Zhao N, et al. Longer-term outcomes of letrozole versus placebo after 5 years of tamoxifen in the NCIC CTG MA.17 trial: analyses adjusting for treatment crossover. J Clin Oncol 2012;30:718-21.

39 Ingle JN, Tu D, Pater JL, et al. Duration of letrozole treatment and outcomes in the placebo-controlled NCIC CTG MA.17 extended adjuvant therapy trial. Breast Cancer Res Treat 2006;99:295-300.

40 Jakesz R, Greil R, Gnant M, et al. Extended Adjuvant Therapy With Anastrozole Among Postmenopausal Breast Cancer Patients: Results From the Randomized Austrian Breast and Colorectal Cancer Study Group Trial 6a. Journal of the National Cancer Institute 2007 Dec 19;99(24):1845-53.

41 Markopoulos C, Dafni U, Misitzis J, et al. Extended adjuvant hormonal therapy with exemestane has no detrimental effect on the lipid profile of postmenopausal breast cancer patients: final results of the ATENA lipid substudy. Breast Cancer Res 2009;11(3):R35.

42 Mamounas EP, Jeong JH, Wickerham DL, et al. Benefit From Exemestane As Extended Adjuvant Therapy After 5 Years of Adjuvant Tamoxifen: Intention-to-Treat Analysis of the National Surgical Adjuvant Breast and Bowel Project B-33 Trial. Journal of Clinical Oncology 2008 Apr 20;26(12):1965-71.

43 Goss P, Mamounas E, Jakesz R, et al. Aromatase Inhibitors (AIs) Versus Not (Placebo/Observation) as Late Extended Adjuvant Therapy for Postmenopausal Women with Early Stage Breast Cancer (BC): Overviews of Randomized Trials of AIs after ~ 5 Years of Tamoxifen. Cancer Research 6924, supplement 3. 15-12-2014. Ref Type: Abstract

44 Burstein HJ, Temin S, Anderson H, et al. Adjuvant Endocrine Therapy for Women With Hormone Receptor-Positive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Focused Update. Journal of Clinical Oncology 2014 May 27.

45 Strasser-Weippl K, Badovinac-Crnjevic T, Fan L, et al. Extended adjuvant endocrine therapy in hormone-receptor positive breast cancer. The Breast22:S171-S175.

46 Bianchini G, Gianni L. An unmet need: tailoring extended adjuvant endocrine therapy. Br J Cancer 2013 Dec 10;109(12):2951-3.

47 Schiavon G, Smith I. Status of adjuvant endocrine therapy for breast cancer. Breast Cancer Research 2014;16(2):206.

48 Smith IE, Yeo B, Schiavon G. The optimal duration and selection of adjuvant endocrine therapy for breast cancer: how long is enough? Am Soc Clin Oncol Educ Book 2014;e16-e24.

49 Mamounas EP, Lembersky B, Jeong JH, et al. NSABP B-42: A Clinical Trial to Determine the Efficacy of Five Years of Letrozole Compared with Placebo in Patients Completing Five Years of Hormonal Therapy Consisting of an Aromatase Inhibitor (AI) or Tamoxifen Followed by an AI in Prolonging Disease-Free Survival in Postmenopausal Women with Hormone Receptor-Positive Breast Cancer. Clinical Breast Cancer 2006 Dec;7(5):416-21.

50 NCT01064635: Letrozole in Treating Breast Cancer in Postmenopausal Women With Stage I, Stage II, or Stage III Breast Cancer Previously Treated With Tamoxifen. http://clinicaltrials.gov/show/NCT01064635 . 5-2-2010. Ref Type: Internet Communication

51 NCT00301457: Different Durations of Adjuvant Anastrozole Therapy After 2 to 3 Years Tamoxifen Therapy in Breast Cancer (DATA). http://clinicaltrials.gov/show/NCT00301457 . 9-3-2006. Ref Type: Internet Communication

52 NCT00295620: Secondary Adjuvant Long Term Study With Arimidex (SALSA). http://clinicaltrials.gov/ct2/show/NCT00295620 . 23-2-2006. Ref Type: Internet Communication

53 Fontein DBY, Nortier JWR, Liefers GJ, et al. High non-compliance in the use of letrozole after 2.5 years of extended adjuvant endocrine therapy. Results from the IDEAL randomized trial. European Journal of Surgical Oncology (EJSO) 2012 Feb;38(2):110-7.

54 International Breast Cancer Study Group. NCT00553410: Letrozole in Preventing Cancer in Postmenopausal Women Who Have Received 4-6 Years of Hormone Therapy for Hormone Receptor-Positive, Lymph Node-Positive, Early-Stage Breast Cancer (SOLE). 2-11-2007. Ref Type: Internet Communication

55 Cardoso F, Piccart-Gebhart M, Van't Veer L, et al. The MINDACT trial: The first prospective clinical validation of a genomic tool. Molecular Oncology 2007 Dec;1(3):246-51.

56 Ma CX, Sanchez CG, Ellis MJ. Predicting endocrine therapy responsiveness in breast cancer. Oncology (Williston Park) 2009 Feb;23(2):133-42.

57 Viale G, Giobbie-Hurder A, Regan MM, et al. Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from breast international group trial 1-98 comparing adjuvant tamoxifen with letrozole. J Clin Oncol 2008;26:5569-75.

58 van de Water W, Fontein DBY, van Nes JGH, et al. Influence of semi-quantitative oestrogen receptor expression on adjuvant endocrine therapy efficacy in ductal and lobular breast cancer: A TEAM study analysis. European Journal of Cancer 2013 Jan;49(2):297-304.

59 Wright C, Nicholson S, Angus B, et al. Relationship between c-erbB-2 protein product expression and response to endocrine therapy in advanced breast cancer. Br J Cancer 1992 Jan;65(1):118-21.

60 Arpino G, Wiechmann L, Osborne CK, et al. Crosstalk between the Estrogen Receptor and the HER Tyrosine Kinase Receptor Family: Molecular Mechanism and Clinical Implications for Endocrine Therapy Resistance. Endocrine Reviews 2008 Apr 1;29(2):217-33.

61 Osborne CK, Shou J, Massarweh S, et al. Crosstalk between Estrogen Receptor and Growth Factor Receptor Pathways as a Cause for Endocrine Therapy Resistance in Breast Cancer. Clinical Cancer Research 2005 Jan 15;11(2):865s-70s.

62 Delpech Y, Wu Y, Hess KR, et al. Ki67 expression in the primary tumor predicts for clinical benefit and time to progression on first-line endocrine therapy in estrogen receptor-positive metastatic breast cancer. Breast Cancer Res Treat 2012 Sep;135(2):619-27.

63 Dowsett M, Smith IE, Ebbs SR, et al. Prognostic Value of Ki67 Expression After Short-Term Presurgical Endocrine Therapy for Primary Breast Cancer. Journal of the National Cancer Institute 2007 Jan 17;99(2):167-70.

64 Cuzick J, Dowsett M, Pineda S, et al. Prognostic Value of a Combined Estrogen Receptor, Progesterone Receptor, Ki-67, and Human Epidermal Growth Factor Receptor 2 Immunohistochemical Score and Comparison With the Genomic Health Recurrence Score in Early Breast Cancer. Journal of Clinical Oncology 2011 Nov 10;29(32):4273-8.

65 Dowsett M, Allred C, Knox J, et al. Relationship Between Quantitative Estrogen and Progesterone Receptor Expression and Human Epidermal Growth Factor Receptor 2 (HER-2) Status With Recurrence in the Arimidex, Tamoxifen, Alone or in Combination Trial. Journal of Clinical Oncology 2008 Mar 1;26(7):1059-65.

66 Barton S, Zabaglo L, A'Hern R, et al. Assessment of the contribution of the IHC4+C score to decision making in clinical practice in early breast cancer. Br J Cancer 2012 May 22;106(11):1760-5.

67 Dowsett M, Salter J, Zabaglo L, et al. Predictive algorithms for adjuvant therapy: TransATAC. Steroids 2011 Jul;76(8):777-80.

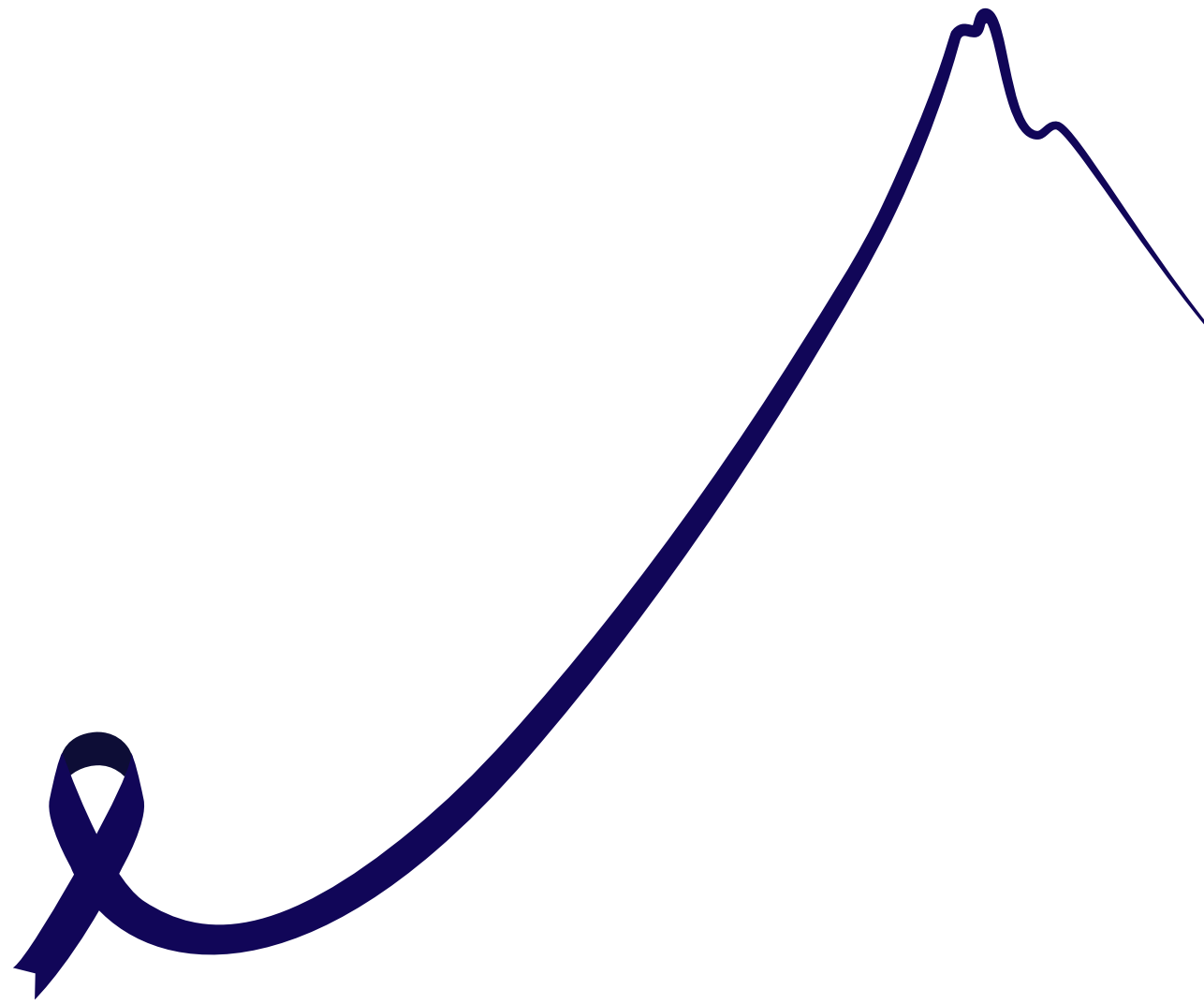
68 Ellis MJ, Tao Y, Luo J, et al. Outcome Prediction for Estrogen Receptor-Positive Breast Cancer Based on Postneoadjuvant Endocrine Therapy Tumor Characteristics. Journal of the National Cancer Institute 2008 Oct 1;100(19):1380-8.

69 Sparano JA. TAILORx: Trial Assigning Individualized Options for Treatment (Rx). Clinical Breast Cancer 2006 Oct;7(4):347-50.

70 Dubsky P, Brase JC, Jakesz R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. Br J Cancer 2013 Dec 10;109(12):2959-64.

71 Sestak I, Cuzick J, Dowsett M, et al. Prediction of Late Distant Recurrence After 5 Years of Endocrine Treatment: A Combined Analysis of Patients From the Austrian Breast and Colorectal Cancer Study Group 8 and Arimidex, Tamoxifen Alone or in Combination Randomized Trials Using the PAM50 Risk of Recurrence Score. Journal of Clinical Oncology 2014 Oct 20.

72 Sestak I, Dowsett M, Zabaglo L, et al. Factors Predicting Late Recurrence for Estrogen Receptor-Positive Breast Cancer. Journal of the National Cancer Institute 2013 Oct 2;105(19):1504-11.

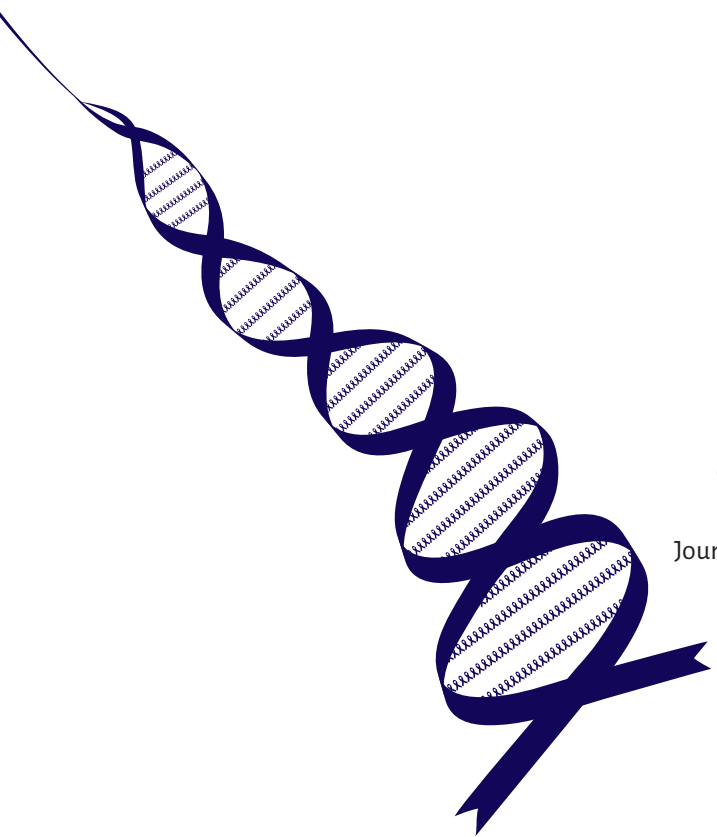


Chapter 3

Optimal duration of extended adjuvant endocrine therapy for early breast cancer; results of the IDEAL-trial (BOOG 2006-05)

E.J. Blok
J.R. Kroep
W.M. Meershoek-Klein Kranenbarg
M. Duijm-de Carpentier
H. Putter
J. van den Bosch
E. Maartense
E. van Leeuwen-Stok
G.J. Liefers
J.W.R. Nortier
E.J.Th. Rutgers
C.J.H. van de Velde
on behalf of the IDEAL Study Group

Journal of the National Cancer Institute
2018; 110(1): 40-48



Abstract

Background: The optimal duration of extended endocrine therapy beyond 5 years after initial aromatase inhibitor based adjuvant therapy for postmenopausal women with hormone-receptor positive breast cancer is still unknown. Therefore, we conducted a clinical trial to compare two different extended endocrine therapy durations.

Methods: In the randomized phase III IDEAL trial, postmenopausal patients with hormone-receptor positive breast cancer were randomly allocated to either 2.5 or 5 years of letrozole after the initial 5 years of any endocrine therapy. The primary endpoint was disease free survival (DFS), and secondary endpoints were overall survival (OS), distant metastasis free interval (DMFi), new primary breast cancer, and safety. Hazard ratios (HRs) were determined using Cox regression analysis. All analyses were by intention to treat principle.

Results: 1824 patients were assigned to either 2.5 years (n=909) or 5 years (n=915) of letrozole, with a median follow-up of 6.6 years. A DFS event occurred in 152 patients in the 5-years group, compared to 163 patients in the 2.5 years group (HR 0.92, 95%CI 0.74-1.16). OS (HR 1.04, 95%CI 0.78-1.38) and DMFi (HR 1.06, 95%CI 0.78-1.45) were not different between both groups. A reduction in occurrence of second primary breast cancer was observed with 5 years treatment (HR 0.39, 95% CI 0.19-0.81). Subgroup analysis did not identify patients that benefit from 5 year extended therapy.

Conclusion: This study showed no superiority of 5 years over 2.5 years of extended adjuvant letrozole, after initial 5 years of adjuvant endocrine therapy.

Introduction

Multiple large clinical trials showed superiority of AI-based adjuvant therapy (either upfront or after 2-3 years of tamoxifen) over 5 years tamoxifen monotherapy. [1-4] Just recently, an EBCTCG meta-analysis showed the superiority of AI monotherapy for 5 years over the sequential therapy of tamoxifen followed by an AI, although the absolute benefit was marginal.[5]

Despite the success of adjuvant endocrine therapy, still 50% of all recurrences occur after the first 5 years, especially in HR-positive breast cancer.[6] Randomized trials showed that 10 years of adjuvant tamoxifen was superior over 5 years, although the benefit on overall survival was not observed.[7-9] The MA.17 study investigated extended endocrine therapy with an AI after 5 years of tamoxifen, by randomly assigning patients to 5 years of letrozole or placebo. At interim-analysis after 2.4 years it was observed that letrozole was superior, leading to early closure and cross-over which hampered the power for long-term follow-up.[10] Although this trial was broadly interpreted as evidence for 5 years therapy extension, the actual evidence before cross-over is only until 2.4 years. The actual benefit of 5 years vs placebo, or the difference in effect between 2.5 and 5 years has never been shown, except for extrapolated subgroup analyses.[10-13]

Until now, all evidence for extended endocrine therapy was obtained in clinical trials that included patients who received tamoxifen monotherapy during the first 5 years of adjuvant therapy. As shown recently in the EBCTCG meta-analysis, adjuvant therapy containing AIs in the first 5 years of adjuvant therapy is superior to tamoxifen monotherapy.[5] However, limited evidence is available for extending AI-based adjuvant therapy beyond 5 years of AI-containing therapy, in particular for the optimal duration of therapy.[14]

We report the results of the phase 3 open label multicenter trial: **I**nvestigation on the **D**uration of **E**xtended **A**djuvant **L**etrozole treatment (IDEAL) trial, which randomly assigned patients to either 2.5 or 5 year letrozole, after receiving any adjuvant endocrine therapy for 5 years. The aim of this trial is to determine the optimal duration of extended endocrine therapy, in particular after receiving AI-based adjuvant therapy.

Materials & Methods

Patients and study design

Postmenopausal women who completed 5 years (± 3 months) of any adjuvant endocrine therapy for early stage hormone-receptor positive (ER and/or PR positive in ≥10% of the nuclei) early breast cancer, were randomized between extending treatment with either 2.5 or 5 years of letrozole (2.5mg daily) (**Figure 1**). Other inclusion criteria were no evidence of breast cancer recurrence at time of randomization, a WHO performance status of 0 or 1, and the initial adjuvant endocrine therapy should be completed for no longer than 2 years. Details on trial design were reported earlier.[15]

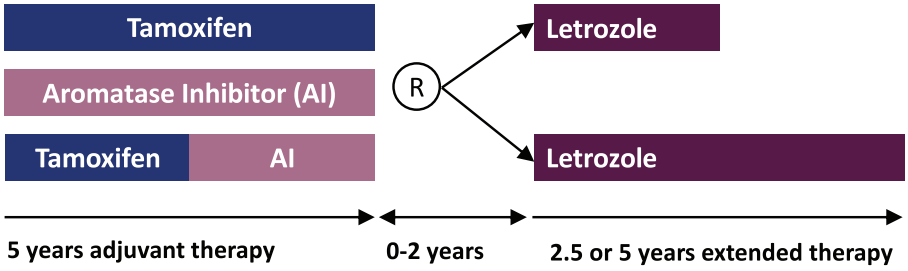


Figure 1: An overview of the trial design.

This study was conducted in 73 hospitals in the Netherlands. Data were collected by the LUMC Datacenter Department of Surgery. The data safety and monitoring board, constituted by an independent statistician, surgeon, and medical oncologist monitored the efficacy endpoints halfway through the trial. Central ethical approval was provided by the ethical committee of the LUMC. All patients provided written informed consent, and were excluded from analysis when consent was withdrawn.

This trial is registered in the Netherlands with the Netherlands Trial Register, NTR3077, the Dutch Breast Cancer Research Group (BOOG 2006-05) and Eudra-CT 2006-003958-16. The study was conducted in compliance with the guidelines of the Declaration of Helsinki, International Conference on Harmonisation and Good Clinical Practice.

Randomization and masking

Randomization was performed by the LUMC Datacenter Department of Surgery in a 1:1 ratio using ALEA software, stratified for prior endocrine therapy regime (5 years

tamoxifen, 5 years AI, or 2-3 years of tamoxifen followed by an AI), time after completion of treatment (0-6 months vs 6-12 months vs 12-24 months), nodal status and the use of adjuvant chemotherapy. All stratification factors were weighted similarly. Pocock's minimisation strategy was used to ensure similar factors in both arms.[16]

Data collection

After providing informed consent, baseline records concerning medical history (including the earlier endocrine therapy), physical examination, mammography, and bone densitometry were collected. Follow-up was conducted annually for at least 5 years with an evaluation of adverse events, disease status, a physical examination, and mammography, with extra visits at 6 and 30 months (latter only for patients in 2.5 year arm to stop allocated therapy).

Endpoints

The primary endpoint of this trial was disease free survival (DFS), defined as the time from randomization to recurrence (either local, regional or distant), new primary breast tumors (DCIS or invasive) or death due to any cause, whichever comes first. Similar to most adjuvant endocrine therapy trials, but in contrast to the definitions defined by Hudis *et al*, second primary non-breast cancer was not included in the definition of DFS.[1, 3, 4, 10, 17] Secondary endpoints were overall survival (OS), distant metastasis free interval (DMFi), new primary breast malignancies (contralateral or new ipsilateral breast cancer), and safety. For safety analysis, adverse events were recorded during active treatment of the patients.

Statistics

It was expected that recurrence rates would be similar in both AI containing arms during the first 2.5 years after randomization, and therefore the power calculations were based on the period after these initial 2.5 years. The objective was to detect an annual decrease of 3.3% in DFS rate in the control arm and 2.0 % in the extended treatment arm (hazard ratio (HR) = 0.60), with a two-sided type I error of 0.05 and power of 80%. Allowing for an annual 2% dropout rate due to loss to follow-up, 126 events, and therefore 1276 patients, were required to detect this difference. Since these 1276 patients needed to be disease free and on treatment after 2.5 years, and with an expected dropout of 30% during the first 2.5 years (due to patients stopping therapy or having a DFS-event in the first 2.5 years after randomization), a number of 1823 patients was required to be randomized.

Despite the fact that the power analysis was performed based on follow-up starting at 2.5 years, it cannot be ruled out that randomization had an influence on either the patient or treating physician during the first 2.5 years since the trial was not blinded. Therefore, all analyses were performed in two parallel ways; the primary analysis starting with all randomized patients on intention-to-treat principle, and the secondary analysis starting at 2.25 years (2.5 years with 10% margin) post-randomization with patients being disease free and on therapy at that time point, after which the treatment arms diverge. Kaplan-Meier analyses were performed for DFS and OS, using stratified log-rank test to determine the level of statistical significance. For DMFi and new primary breast malignancies cumulative incidence curves were estimated, accounting for death as competing risk. Furthermore, for all endpoints, univariate stratified Cox regression analysis was used to determine the hazard ratio (HR). The proportional hazards assumption for treatment (the only variable for which proportional hazards was assumed) was checked using Schoenfeld residuals. Stratified Cox regression within subgroups was used to perform subgroup analysis. For analyses of the adverse events, chi-square tests were used to assess which AE occurs more frequently in which treatment arm, applying Bonferroni correction to correct for multiple testing. All analyses were performed using SPSS 23.0, data visualization was performed using GraphPad Prism 6.05 and R 3.2.2.

All statistical tests were two-sided and a P-value of less than 0.05 was considered statistically significant.

Results

Study population

As planned, 1824 patients were randomized between April 2007 and November 2011 in 73 participating hospitals in the Netherlands (909 patients in 2.5 years group, 915 patients in 5 years group). The median follow-up of these patients was 6.6 year (inter quartile range (IQR) 5.3-7.5 years). Of these, 3 patients withdrew their consent and were excluded for the primary analysis starting at randomization, leaving 908 patients in the 2.5 years group and 913 patients in the 5 years group (**Figure 2**). All other patients were included in the intention-to-treat analysis. Furthermore, 482 patients encountered a DFS event or stopped with therapy before they reached 2.25 year, leaving 1339 patients for the secondary analysis after 2.25 years. In this secondary analysis, the median follow-up was 6.6 years (IQR 5.2-7.5 years)

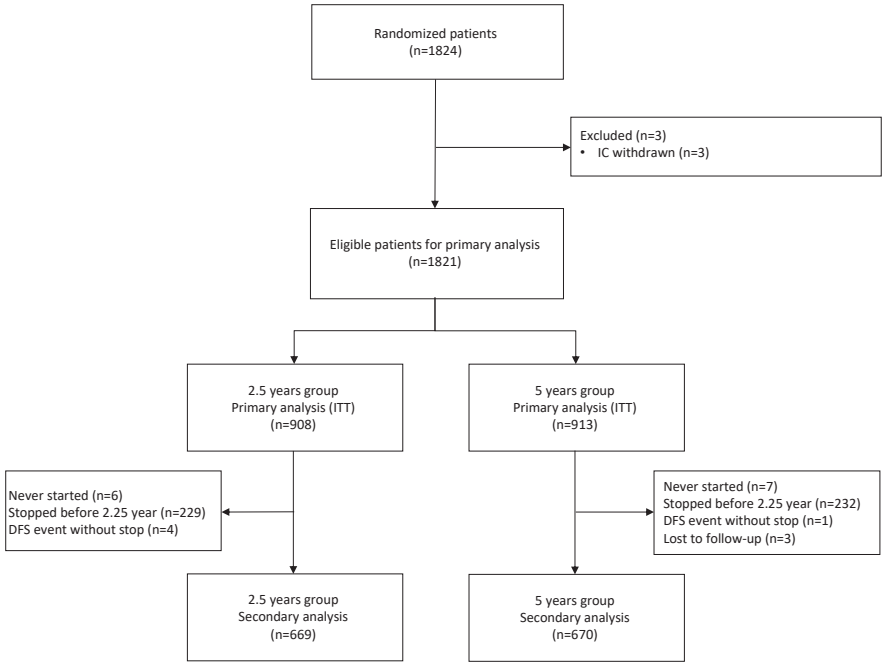


Figure 2: A consort diagram showing the flowchart of the trial.
N:number, IC: informed consent, ITT: intention-to-treat, DFS: disease-free survival

Baseline characteristics for the randomized eligible patients are shown in **Table 1**. There were no statistically significant differences observed between both arms. The majority of patients received AI-based adjuvant therapy, either upfront (28.8%) or after 2-3 years of tamoxifen (59.0%). Only 12.2 percent were AI-naïve and received 5 years of tamoxifen. Most patients (88.6%) continued with extended therapy within 6 months after regular adjuvant endocrine therapy.

Compliance

To assess the capacity of patients to endure extended endocrine therapy, compliance was monitored closely in this trial. A total of 629 patients stopped therapy earlier than planned (34.6%). In the group allocated to 2.5 years, 241 (26.5%) patients stopped early, for which the main reasons were symptoms or adverse events (n=156), a study endpoint (recurrence, new primary tumor or death) (n=30), and treatment refusal (n=24). In the 5-years group 388 patients (42.5%) stopped before 5 years of treatment, for which the main reasons were symptoms or adverse events (n=212), a study endpoint (recurrence, new primary tumor or death) (n=78), and treatment refusal (n=46) (**Figure 3**). Furthermore, 104 patients continued with therapy beyond their allocated treatment

duration with a median overtreatment of 4 months, 13 patients never started therapy and 3 patients withdrew consent, limiting the total compliance to 59.9%.

Table 1. Baseline clinicopathological features of all randomized patients per treatment arm

Subgroups	Treatment Arm			
	2.5 years letrozole		5 years letrozole	
	N	(%)	N	(%)
Age at randomization, y				
<55	250	(27.5%)	260	(28.5%)
55-65	386	(42.5%)	375	(41.1%)
65-75	210	(23.1%)	201	(22.0%)
>75	62	(6.8%)	77	(8.4%)
Nodal status				
pN0	227	(25.0%)	223	(24.4%)
pN0(i+)	10	(1.1%)	12	(1.3%)
pN1(mi)	105	(11.6%)	105	(11.5%)
pN1: 1-3 pos	433	(47.7%)	431	(47.2%)
pN2: 4-9 pos	97	(10.7%)	104	(11.4%)
pN3: >=10 pos	30	(3.3%)	29	(3.2%)
Tumor type				
ductal	683	(75.2%)	732	(80.2%)
mucinous	9	(1.0%)	7	(0.8%)
medullar	3	(0.3%)	4	(0.4%)
lobular	165	(18.2%)	131	(14.3%)
other	47	(5.2%)	39	(4.3%)
Histological grade				
grade 1	156	(17.2%)	130	(14.2%)
grade 2	380	(41.9%)	394	(43.2%)
grade 3	270	(29.7%)	296	(32.4%)
unknown	102	(11.3%)	93	(10.1%)
Progesterone receptor status				
Negative	160	(17.6%)	182	(19.9%)
Positive ≥10%	712	(78.4%)	697	(76.3%)
HER2 status				
0	193	(45.7%)	199	(47.0%)
1+	95	(22.5%)	93	(22.0%)
2+	47	(11.1%)	51	(12.1%)
3+	81	(19.2%)	78	(18.4%)
Performed final surgery				
breast conserving	445	(49.0%)	443	(48.5%)
mastectomy	460	(50.7%)	468	(51.3%)
Prior chemotherapy				
no	291	(32.0%)	287	(31.4%)
yes	617	(68.0%)	626	(68.6%)

Table 1. continued

Subgroups	Treatment Arm			
	2.5 years letrozole		5 years letrozole	
	N	(%)	N	(%)
Prior endocrine treatment				
5 years tamoxifen	109	(12.0%)	113	(12.4%)
5 years AI	261	(28.7%)	263	(28.8%)
2-3 years tam-> 3-2 years AI	538	(59.3%)	537	(58.8%)
Time after stop hormonal therapy, mos				
0 to <6	803	(88.4%)	811	(88.8%)
6 to <12	48	(5.3%)	47	(5.1%)
12-27	57	(6.3%)	55	(6.0%)

Endpoints

At the moment of database lock (December 22th, 2016), 315 out of 1821 patients in the primary analysis had encountered a DFS event, of which 163/908 (18.0%) in the 2.5 year arm and 152/913 (16.6%) in the 5 years arm (**Table 2**). The hazard ratio (HR) for DFS was 0.92 (95% CI 0.74-1.16, Log-rank P=0.49) for patients in the 5 year group, compared to the 2.5 year group (**Figure 4A**). A preplanned subgroup analysis showed that there is no individual subgroup which benefits statistically significant from extended adjuvant endocrine therapy up to 5 year (**Figure 5**). The proportional hazards assumption for treatment was not found to be violated.

Furthermore, no statistically significant effect on either overall survival (**Figure 4B**) or distant recurrences (**Figure 4C**) was shown with respective HRs of 1.04 (OS, 95% CI 0.78-1.38, Log-rank P=0.79) and 1.06 (DMFi, 95% CI 0.78-1.45, Log-rank P=0.71). For second primary breast malignancies (**Figure 4D**), 27 (3.1%) events were observed in the 2.5-year group and 10 (1.1%) in the 5-year group, which was statistically significant (HR 0.39, 95% 0.19-0.81, Log-rank P=0.01).

In the secondary analysis (**Figure 6**), in which patients who encountered an event or stopped therapy before 2.25 years were excluded, 86 DFS events were observed during follow-up in the 2.5 year arm, and 74 events in the 5 year arm (HR 0.88, 95% CI 0.64-1.21) (**Table 2**). Of these events, 15 second primary breast malignancies were observed in the 2.5 year arm, and 6 in the 5 year arm (HR 0.42, 95% CI 0.16-1.11).

Table 2. An overview of the number of events in both arms and the subsequent hazard ratio (HR), both for the primary population, and the secondary population who were disease free and on therapy at 2.25 years*

Endpoints	Treatment arm		HR (95% CI)
	5 year letrozole	2.5 year letrozole	
	No. of events	No. of events	
DFS (full population)	152/913	163/908	0.92 (0.74-1.16)
local recurrence	14	12	1.06 (0.49-2.31)
regional recurrence	14	10	1.27 (0.55-2.92)
distant recurrence	86	78	1.06 (0.78-1.45)
2nd primary breast cancer	10	27	0.39 (0.19-0.81)
death any cause	104	96	1.04 (0.78-1.38)
DFS (after 2.25 year)	74	86	0.88 (0.64-1.21)
local recurrence	10	8	1.17 (0.46-2.98)
regional recurrence	6	7	0.92 (0.30-2.76)
distant recurrence	35	47	0.75 (0.48-1.17)
2nd primary breast cancer	6	15	0.42 (0.16-1.11)
death any cause	45	40	1.06 (0.68-1.65)

*CI=confidence interval; DFS=Disease free survival

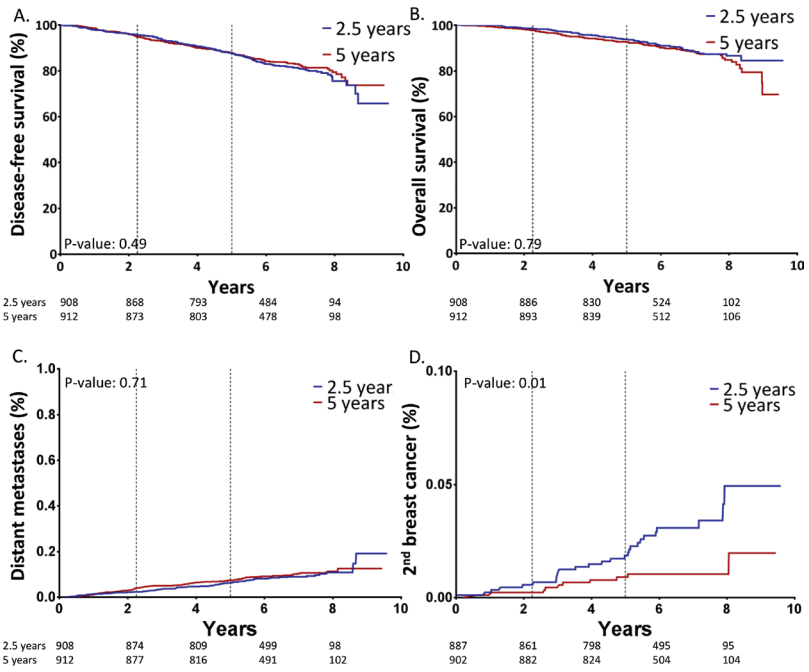


Figure 4: Kaplan Meier analysis. Results are shown for (A) disease free survival (DFS), (B) overall survival (OS), (C) distant metastasis free interval (DMFi), and (D) new primary breast cancer, including all randomized patients based on intention-to-treat principle. Log-rank tests were used to assess the differences between groups within each graph (reported as p-value).

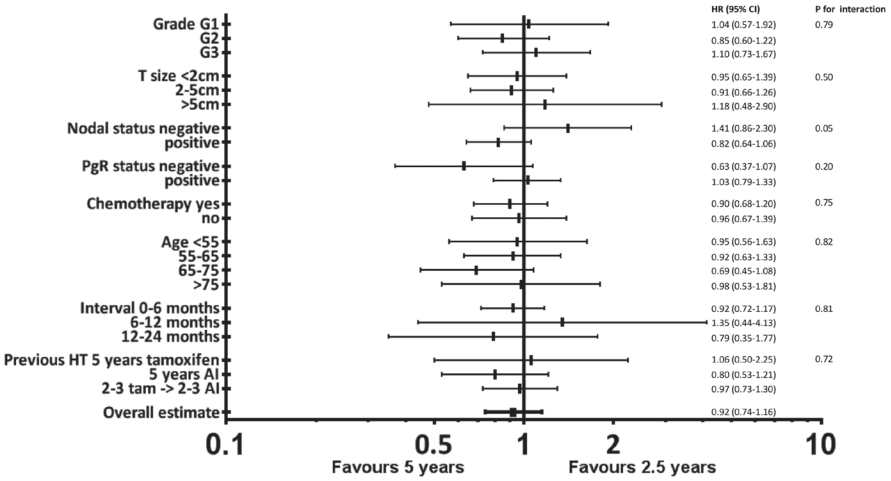


Figure 5: A pre-planned subgroup analysis. All values were determined using two-sided Cox-regression analysis. Error bars represent 95% confidence intervals.
HR: hazard ratio, CI: confidence interval, T size: tumor size, PgR: progesterone receptor, HT: hormonal therapy, AI: aromatase inhibitor, tam: tamoxifen

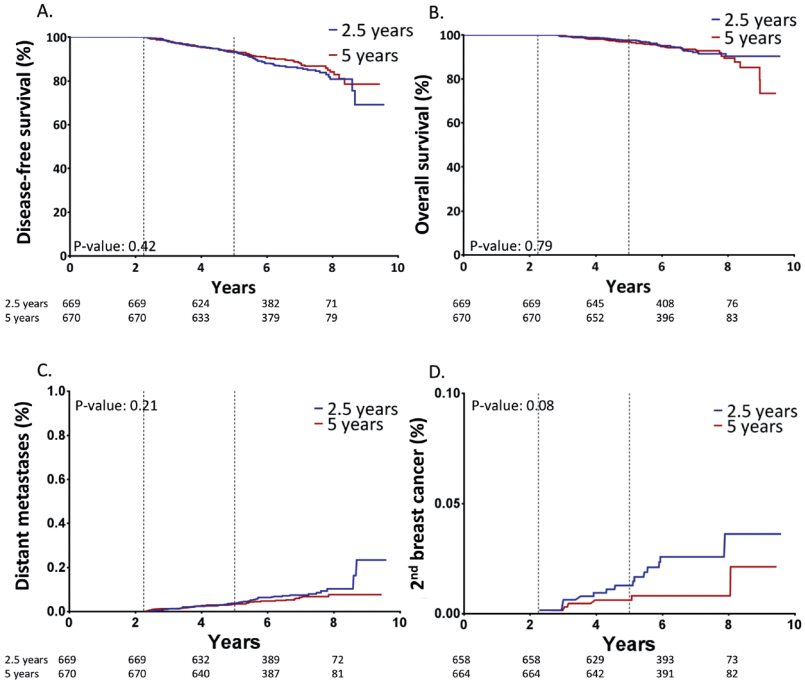


Figure 6: Secondary analysis. Results are shown for (A) disease free survival (DFS), (B) overall survival (OS), (C) distant metastasis free interval (DMFi), and (D) new primary breast cancer, including all patients that were disease free and on therapy at 2.25 years. Log-rank tests were used to assess the differences between groups within each graph (reported as p-value).

Table 3. An overview of the most frequently reported adverse events, stratified per grade and treatment arm*

Adverse events	2.5 year letrozole					5 year letrozole					Total (%)
	Grade 1	Grade 2	Grade 3	Grade 4	Any grade, No. (%)	Grade 1	Grade 2	Grade 3	Grade 4	Any grade, No. (%)	
Arthralgia	72	40	7	0	119 (13.2%)	70	48	13	2	133 (14.7%)	252 (14.0%)
Hot flashes	67	24	5	0	96 (10.5%)	69	40	6	3	118 (13.1%)	214 (11.8%)
Osteoporosis	39	26	3	0	68 (7.5%)	61	54	1	0	116 (12.7%)	184 (10.2%)
Fatigue	46	17	5	0	68 (7.5%)	50	34	3	1	88 (9.7%)	156 (8.6%)
Joint range of motion decreased	43	14	2	0	59 (6.5%)	33	21	2	0	56 (6.2%)	115 (6.4%)
Alopecia	51	6	2	0	59 (6.5%)	45	7	1	1	54 (6.0%)	113 (6.3%)
Depression	34	18	5	0	57 (6.2%)	23	20	4	0	47 (5.2%)	104 (5.8%)
Back pain	30	20	5	0	55 (6.1%)	19	22	2	2	45 (5.0%)	100 (5.5%)
Fracture	2	17	5	1	25 (2.8%)	6	24	14	1	45 (5.0%)	70 (3.9%)
Total	935	496	126	15	1580 (70.1%)	983	681	155	34	1860 (71.8%)	3440 (71.4%)

* All events with a frequency over 5% in one of the arms are shown.

Safety

In all patients who started therapy (n=1806), 3440 adverse events were reported by 1289 patients. Of these events, 1580 were reported by 640 (70.1%) patients in the 2.5 year arm during active treatment, and 1860 were reported by 649 patients (71.8%) in the 5 year arm during treatment. Of all events, 90.3% was graded as 1 or 2, and there was no difference in the proportion of grade 3/4 events between both groups (2.5yr: 8.8%, 5yr: 10.0%, X^2 p=0.43) (data not shown).

A total of 368 patients stopped therapy due to AEs, of which 156 in the 2.5-years arm (17.3%) and 212 in the 5-years arm (23.5%) In patients allocated to 5 years of therapy, the majority of events (n=1481, 79.6%) occurred during the first 2.5 years. In total, 85.8% of the patients (n=182) in the 5 years group that ceased therapy due to side effects, did this before 2.5 years. The frequency of adverse events is reported in **Table 3**, in which all events with a frequency over 5% in one of the arms are shown. Most frequently reported AEs were arthralgia, reported by 252 patients (14.0%) , hot flashes (n=214, 11.8%) and osteoporosis (n=184, 10.2%). The most reported grade 3/4 AEs were arthralgia (n=22) and fractures (n=21).

Discussion

This study has shown that, after receiving any adjuvant endocrine therapy for 5 years, there is no statistically significant difference in disease related outcomes between patients treated with either 2.5 or 5 years of letrozole at a median follow-up of 6.6 years, with the exception of preventing new primary breast malignancies. Subgroup analysis showed that there was no benefit of 5 years of extended therapy regarding DFS for any specific subgroup. Furthermore, no interaction between subgroups was observed.

Additionally, we observed a statistically significant decrease in second primary breast malignancies in patients treated with 5 years of extended therapy. This observation was in agreement with the MA.17R trial, in which most of the effect of 5 years letrozole after 10 years of earlier therapy was accounted to prevention of contralateral breast cancer.[18] It could be argued that extended endocrine adjuvant therapy with aromatase inhibitors beyond 7.5 years is secondary prevention rather than actual adjuvant therapy preventing relapse of the earlier breast cancer. This preventive effect

has already been shown in multiple clinical trials in healthy women without breast cancer, using both tamoxifen and AIs.[19-25]

This study did not question whether AI-containing adjuvant therapy should be extended beyond the first 5 years. The MA.17 and MA.17R trials already showed that 5 years of letrozole was superior to placebo after the initial 5 years of tamoxifen monotherapy, and that a further extension up to 10 years of AIs led to a further improvement in DFS.[13, 18] However, death from any cause was not included in their definition of DFS, and the statistically significant effect on DFS in MA.17R was mainly attributed to a decrease in second primary breast cancers.[18] Furthermore, the results of both MA.17 and MA.17R are not valid for the majority of patients, who nowadays receive upfront AI as adjuvant endocrine therapy.[26]

The B42 trial, presented recently at SABCS 2016, compared 5 years of letrozole to placebo after initial AI-containing adjuvant therapy. They did not show a benefit on DFS in the overall patient group and subgroups [27]The DATA trial, presented at the same conference, showed that there is no statistically significant benefit of 6 years anastrozole over 3 years anastrozole, after initial 2-3 years of tamoxifen.[28] In contrast to the B42 trial and our results, their subgroup analysis suggested a statistically significant benefit for higher risk patients (node positive, tumor size larger than pT2) and for tumors expressing both ER and PR.

Combining these recent results, there is no evidence for therapy extension for the general hormone receptor positive postmenopausal breast cancer patient after an AI in the first 5 years. Data on high-risk subgroups, reflected by tumor size, nodal status, or hormone receptor subgroups are discordant. It is unclear why, in general, there is a lack of extended therapy effect in the population that received AIs earlier. A possible explanation could be the relative inferiority of tamoxifen during the first 5 years, which leaves a possibility for benefit of extended therapy. A second explanation might be therapy resistance. In metastatic disease, it is well known that mutations in the gene encoding for ER, are associated to resistance against AIs.[29, 30] Although this has not been studied, a similar mechanism could play a role in dormant tumor cells, making them resistant against the adjuvant treatment and causing the extended therapy to have no additional benefit.

A number of clinical trials studying the extension of AI-based adjuvant therapy are still ongoing.[14] In case future studies will show a benefit of extended AI adjuvant therapy, the results of this trial show that the effect is limited to 7.5 years of total treatment duration. However, it cannot be ruled out that there is an effect in a subgroup of patients. For this, future explorative subgroup analyses will be performed, and follow-up will be extended up to 10 years. Furthermore, a translational side study is initiated, to explore biomarkers capable of predicting extended therapy benefit.

The rate of patients reporting AEs is similar in both arms, although the absolute count of AEs is higher in the 5 year group. However, since adverse events were only recorded during active treatment, the frequency of AEs in the 2.5 years group might be underreported since there was no registration of side effects in the second 2.5 years in which there was no therapy. The frequency of specific adverse events, like e.g. hot flashes, is lower than expected based on earlier studies. In the MA.17 trial, 5 year of letrozole was associated to 47% of patients reporting hot flashes, whereas in this trial only 12% of patients reported these symptoms.[31] Most likely, these differences are due to differences in trial design. In the MA.17 trial, all patients where AI-naïve, whereas 88% patients in this trial had earlier received treatment with an AI and were therefore less likely to report the side effects. Furthermore, selection bias might have occurred, since patients that experienced side effects during regular adjuvant therapy, would have been less likely to participate in this trial.

A limitation of this trial is the upfront randomization. After randomization, there was approximately 30% drop-out before the moment that the treatment arms actually diverged, which could have led to additional random differences between both arms. However, this drop-out was accounted for in the sample size calculation, and therefore did not influence the statistical power of the analyses. A second limitation is the open-label design. In combination with the upfront randomization, this could have influenced the patient or clinician in their decisions. However, drop-out was similar in both groups during the first 2.5 years, although a small bias cannot be excluded. In order to prevent an attrition bias during the first 2.5 years, the primary analysis started at randomization and not at the moment that the treatment arms diverged.

In summary, we have shown that the effect on any disease-related outcomes of 5 years of extended letrozole was not superior over 2.5 years of extended therapy with

letrozole, after 5 years of any regular adjuvant endocrine therapy, except for a small difference in the occurrence of new primary breast malignancies. Although this study did not show the added value of extended use of AI-containing adjuvant therapy in itself, it has shown that whenever extended AI-containing adjuvant therapy is considered, extended therapy longer than 2.5 years will not lead to a further reduction in disease free or overall survival.

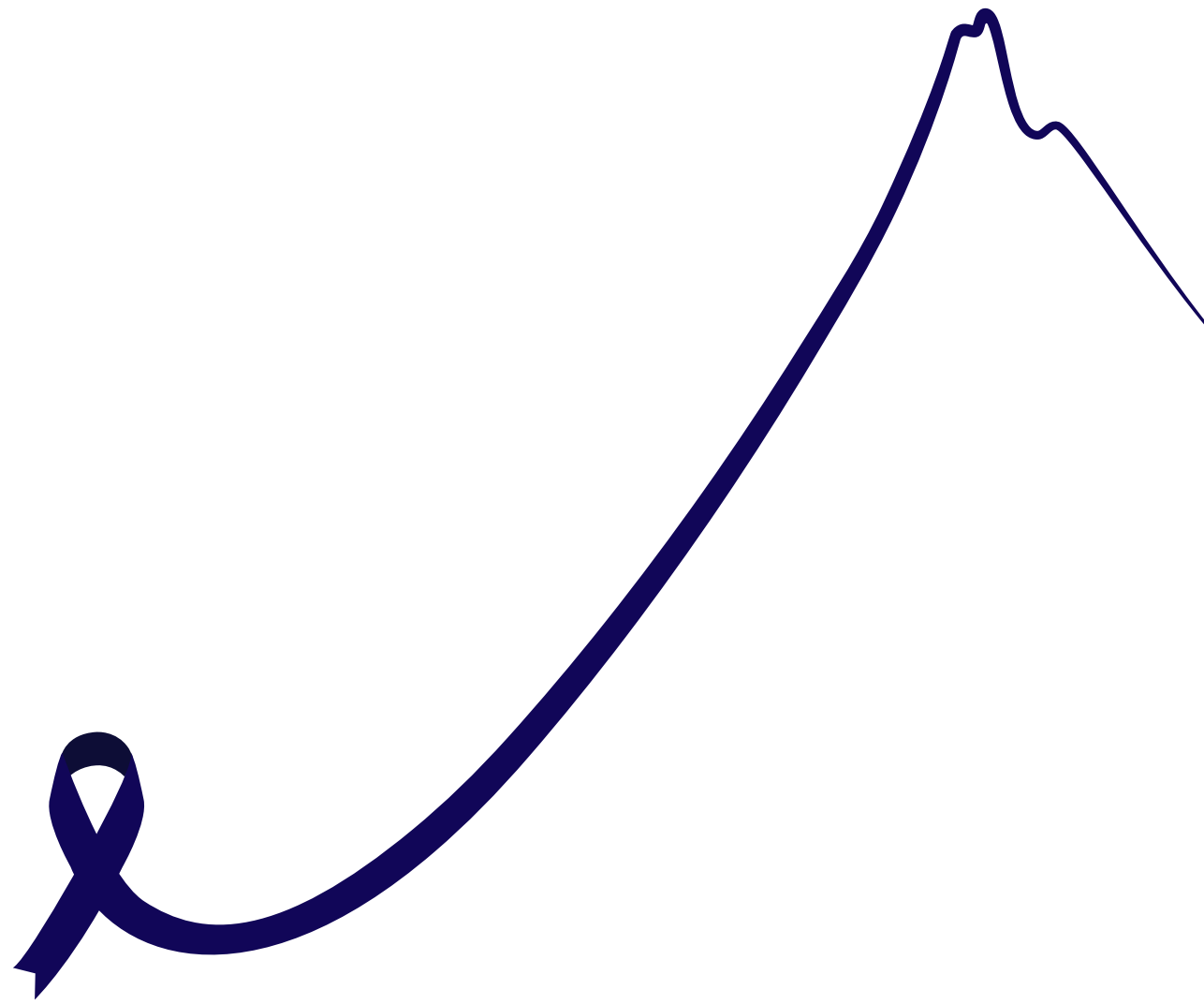
Funding

This work was supported by Novartis by means of an independent educational grant (CFEM345DNL03).

References

1. van de Velde CJ, Rea D, Seynaeve C et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-331.
2. Thurlimann B, Keshaviah A, Coates AS et al. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med 2005;353:2747-2757.
3. Howell A, Cuzick J, Baum M et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. The Lancet 2005;365(9453):60-62.
4. Coombes RC, Hall E, Gibson LJ et al. A Randomized Trial of Exemestane after Two to Three Years of Tamoxifen Therapy in Postmenopausal Women with Primary Breast Cancer. N Engl J Med 2004;350(11):1081-1092.
5. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
6. Colleoni M, Sun Z, Price KN et al. Annual Hazard Rates of Recurrence for Breast Cancer During 24 Years of Follow-Up: Results From the International Breast Cancer Study Group Trials I to V. JCO 2016;34(9):927-935.
7. Al-Mubarak M, Tibau A, Templeton AJ et al. Extended Adjuvant Tamoxifen for Early Breast Cancer: A Meta-Analysis. PLoS ONE 2014;9(2):e88238.
8. Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).
9. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381:805-816.
10. Goss PE, Ingle JN, Martino S et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med 2003;349:1793-1802.
11. Jin H, Tu D, Zhao N et al. Longer-Term Outcomes of Letrozole Versus Placebo After 5 Years of Tamoxifen in the NCIC CTG MA.17 Trial: Analyses Adjusting for Treatment Crossover. JCO 2012;30(7):718-721.
12. Ingle JN, Tu D, Pater JL et al. Intent-to-treat analysis of the placebo-controlled trial of letrozole for extended adjuvant therapy in early breast cancer: NCIC CTG MA.17. Ann Oncol 2008;19:877-882.
13. Goss PE, Ingle JN, Martino S et al. Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17. J Natl Cancer Inst 2005;97:1262-1271.
14. Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.
15. Fontein DBY, Nortier JWR, Liefers GJ et al. High non-compliance in the use of letrozole after 2.5 years of extended adjuvant endocrine therapy. Results from the IDEAL randomized trial. European Journal of Surgical Oncology (EJSO) 2012;38(2):110-117.
16. Pocock SJ, Simon R. Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial. 1975;31(1):103-115.
17. Hudis CA, Barlow WE, Costantino JP et al. Proposal for Standardized Definitions for Efficacy End Points in Adjuvant Breast Cancer Trials: The STEEP System. JCO 2007;25(15):2127-2132.
18. Goss PE, Ingle JN, Pritchard KI et al. Extending Aromatase-Inhibitor Adjuvant Therapy to 10 Years. N Engl J Med 2016;375(3):209-219.
19. Cuzick J, Sestak I, Cawthorn S et al. Tamoxifen for prevention of breast cancer: extended long-term follow-up of the IBIS-I breast cancer prevention trial. The Lancet Oncology16(1):67-75.
20. Fisher B, Costantino JP, Wickerham DL et al. Tamoxifen for the Prevention of Breast Cancer: Current Status of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. Journal of the National Cancer Institute 2005;97(22):1652-1662.
21. Cummings SR, Eckert S, Krueger KA. The effect of raloxifene on risk of breast cancer in postmenopausal women: Results from the more randomized trial. JAMA 1999;281(23):2189-2197.
22. Cuzick J, Sestak I, Forbes JF et al. Anastrozole for prevention of breast cancer in high-risk postmenopausal women (IBIS-II): an international, double-blind, randomised placebo-controlled trial. The Lancet 1922;383(9922):1041-1048.

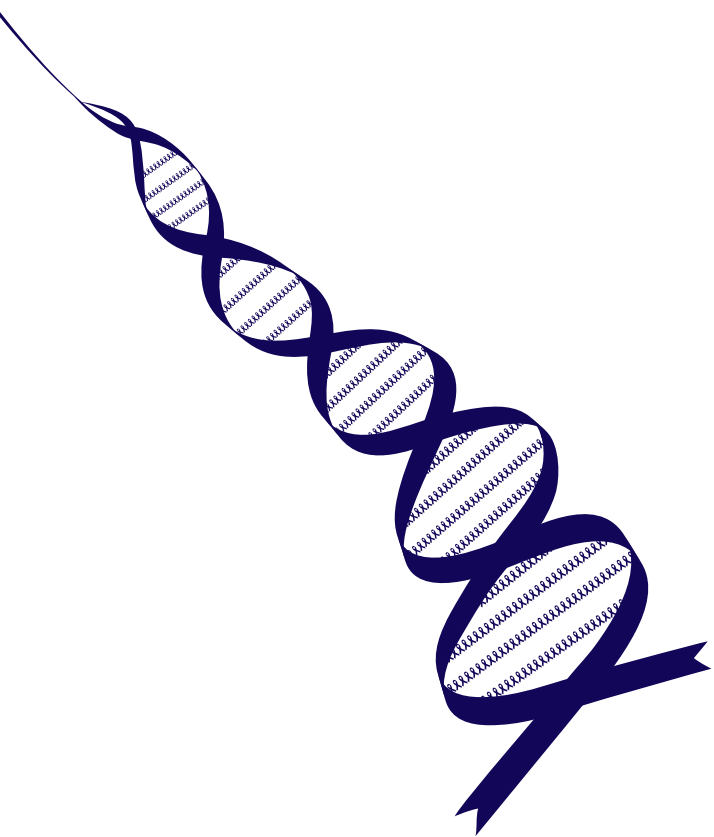
23. Goss PE, Ingle JN, Al+¬s-Mart+¡nez JE et al. Exemestane for Breast-Cancer Prevention in Postmenopausal Women. N Engl J Med 2011;364(25):2381-2391.
24. Howell A, Anderson AS, Clarke RB et al. Risk determination and prevention of breast cancer. Breast Cancer Research 2014;16(5):1-19.
25. Rahman RL, Pruthi S. Chemoprevention of Breast Cancer: The Paradox of Evidence versus Advocacy Inaction. Cancers (Basel) 2012;4(4):1146-1160.
26. Cuzick J. Statistical controversies in clinical research: long-term follow-up of clinical trials in cancer. Annals of Oncology 2015.
27. Mamounas EP, Bandos H, Lembersky BC et al. Abstract S1-05: A randomized, double-blinded, placebo-controlled clinical trial of extended adjuvant endocrine therapy (tx) with letrozole (L) in postmenopausal women with hormone-receptor (+) breast cancer (BC) who have completed previous adjuvant tx with an aromatase inhibitor (AI): Results from NRG Oncology/NSABP B-42. Cancer Research 2017;77(4 Supplement):S1-05.
28. Tjan-Heijnen VC, Van Hellemond IE, Peer PG et al. Abstract S1-03: First results from the multicenter phase III DATA study comparing 3 versus 6 years of anastrozole after 2-3 years of tamoxifen in postmenopausal women with hormone receptor-positive early breast cancer. Cancer Research 2017;77(4 Supplement):S1-03.
29. Jeselsohn R, Yelensky R, Buchwalter G et al. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. Clin Cancer Res 2014;20(7):1757-1767.
30. Jeselsohn R, Buchwalter G, De Angelis C et al. ESR1 mutations: a mechanism for acquired endocrine resistance in breast cancer. Nat Rev Clin Oncol 2015;12(10):573-583.
31. Goss PE, Ingle JN, Martino S et al. A Randomized Trial of Letrozole in Postmenopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. N Engl J Med 2003;349(19):1793-1802.



Chapter 4

Relevant factors for the optimal duration of extended endocrine therapy in early breast cancer

E.J. Blok
J.R. Kroep
W.M. Meershoek-Klein Kranenbarg
M. Duijm-de Carpentier
H. Putter
G.J. Liefers
J.W.R. Nortier
E.J.Th. Rutgers
C.S. Seynaeve
C.J.H. van de Velde



Breast Cancer Res Treat.
2018 Apr;168(2):413-420

Abstract

For postmenopausal patients with hormone-receptor positive early breast cancer, the optimal subgroup and duration of extended endocrine therapy is not clear yet. The aim of this study using the IDEAL patient cohort, was to identify a subgroup for which longer (5 years) extended therapy is beneficial over shorter (2.5 years) extended endocrine therapy.

In the IDEAL trial, 1824 patients who completed 5 years of adjuvant endocrine therapy (either 5 years of tamoxifen (12%), 5 years of an AI (29%) or a sequential strategy of both (59%)), were randomized between either 2.5 or 5 years of extended letrozole. For each prior therapy subgroup, the value of longer therapy was assessed for both node-negative and node-positive patients using Kaplan Meier and Cox regression survival analyses.

In node-positive patients, there was a significant benefit of 5 years (over 2.5 years) of extended therapy (disease-free survival (DFS) HR 0.67, p=0.03, 95% CI 0.47-0.96). This effect was only observed in patients who were treated initially with a sequential scheme (DFS HR 0.60, p=0.03, 95% CI 0.38-0.95). In all other subgroups, there was no significant benefit of longer extended therapy. Similar results were found in patients who were randomized for their initial adjuvant therapy in the TEAM trial (DFS HR 0.37, p=0.07, 95% CI 0.13-1.06), although this additional analysis was underpowered for definite conclusions.

This study suggests that node-positive patients could benefit from longer extended endocrine therapy, although this effect appears isolated to patients treated with sequential endocrine therapy during the first 5 years and needs validation and long-term follow-up.

Introduction

In hormone receptor-positive (HR+) breast cancer, adjuvant endocrine therapy is used to decrease the risk for recurrence, and improve the overall survival (OS). Where tamoxifen for five years has been the standard adjuvant endocrine therapy for a long period of time, currently, treatment regimens for adjuvant endocrine therapy are mostly based on 5 years of an aromatase inhibitor (AI), or a sequential strategy of tamoxifen followed by an AI. Among others, the Tamoxifen Exemestane Adjuvant Multinational (TEAM) trial showed that after 5 and 10 years of follow-up, there was no difference in disease-free survival (DFS) between patients randomized to either tamoxifen followed by exemestane, or exemestane monotherapy.[1, 2] These results were confirmed in a meta-analysis performed by the Early Breast Cancer Trialists Collaborative Group (EBCTCG).[3]

Despite the value of adjuvant endocrine therapy, it is known that the risk for recurrence in HR+ breast cancer remains linear up to at least 15 years after diagnosis prompting to study the value of extended endocrine therapy. After 5 years of tamoxifen, it has been established that extended therapy beyond 5 years leads to a modest reduction in recurrences, but not in overall survival.[4, 5] This has been particularly observed for patients with node-positive disease.[6]

The value of extended endocrine therapy after a 5-years regimen including an AI (either upfront or after 2-3 years of tamoxifen) is less clear. Recently, three independent studies did not show a significant benefit of (longer) extended endocrine therapy for the total study population.[7-9] In the NSABP B-42 trial, patients who earlier received either 5 years of an AI, or a sequential treatment of tamoxifen followed by an AI until 5 years, were randomized between 5 years of extended letrozole, or placebo. After 5 years, there was no significant benefit of 5 years of letrozole over placebo. In the subgroup analysis however, a significant benefit for patients who received prior tamoxifen followed by an AI (HR 0.75, p=0.04) was found, which was not observed in patients who were treated upfront with AI monotherapy for 5 years (HR 0.91, p=0.34).[7]

In the Dutch 'Investigation on the Duration of Extended Adjuvant Letrozole treatment' (IDEAL) trial, 1824 postmenopausal patients who received any form of primary adjuvant endocrine therapy for 5 years, were randomized between extended letrozole for 2.5 or 5 years. The results of this trial were published recently by our group, and

identified no subgroup that benefitted significantly from 5 instead of 2.5 years of extended therapy.[9] In the IDEAL trial, approximately 60% was treated initially with the sequential scheme, whereas 30% was treated with an upfront aromatase inhibitor only, and approximately 10% was treated with tamoxifen monotherapy.

In the Dutch study on 'Duration of Anastrozole therapy after two to three years Tamoxifen as Adjuvant therapy' (DATA), postmenopausal patients were randomized after 2-3 years of tamoxifen between 3 years of anastrozole (standard arm, duration endocrine therapy 5-6 years in total) or 6 years of anastrozole (extended duration, 8-9 years in total). Also in this trial, no effect of extended AI (anastrozole) therapy was shown for the total population. However, this study did observe a significant benefit of longer AI therapy in high-risk subgroups, in particular patients with lymph-node positive disease.[8]

Combining the conclusions on the subgroup analyses of the NSABP B-42 and DATA trials, it is suggested that extended therapy might be the most beneficial for node-positive patients who were previously treated with tamoxifen followed by an AI. However, the optimal duration of extended therapy is not clear, since the regimens and populations in both trials differ too much for direct comparisons. In view of the above mentioned data, we performed an additional subgroup analysis in the IDEAL trial. The aims of the current subanalyses were to investigate the effect of primary adjuvant treatment and nodal status on the optimal duration of extended adjuvant endocrine therapy. Furthermore, similar analyses were conducted in the subgroup of patients that previously participated in the TEAM trial, as this subgroup was randomized for the initial therapy.

Methods

IDEAL trial cohort

In the phase 3 IDEAL trial, 1824 postmenopausal patients were randomized between 2.5 or 5 years of letrozole, after 5 years of any type of adjuvant endocrine therapy for early HR+ breast cancer. Patients needed to be disease-free at the moment of randomization. Furthermore, a maximum of 2 years was allowed between finishing earlier endocrine therapy and starting extended therapy. As the treatment arms during the first 2.5 years were equal, no differences can be expected during this period. Therefore for the current analysis, patients that encountered an event or stopped therapy during the first 2.5 years were excluded, and the survival analysis

started at 2.5 years after randomization at which time point the treatment arms diverge. Details of the trial, data collect and the primary results have recently been reported elsewhere.[9,10]

A total of 438 IDEAL patients (24%) also participated in the TEAM-trial during the first 5 years of their adjuvant endocrine therapy. In that phase III study, postmenopausal patients with early HR+ breast cancer were randomized at diagnosis between 5 years of exemestane, or 2.5 years of tamoxifen followed by 2.5 years of exemestane (sequential scheme). In case they were disease-free and finished 5 years of therapy, and their hospital participated in the IDEAL trial, they were eligible for inclusion in the IDEAL trial. In order to correct for a possible allocation bias in the distribution of previous endocrine therapy between node-negative and node-positive patients, all analysis were repeated in the cohort of patients that participated in the TEAM trial as these patients were not subjected to allocation bias due to the randomization already at primary diagnosis.

The IDEAL trial is registered in the Netherlands with the Netherlands Trial Register, NTR3077, the Dutch Breast Cancer Research Group (BOOG 2006-05) and Eudra-CT 2006-003958-16. The original study was conducted in compliance with the guidelines of the Declaration of Helsinki, International Conference on Harmonisation and Good Clinical Practice.

Endpoints

The primary endpoint of the IDEAL trial was disease-free survival (DFS) defined as the time from randomization to recurrence (either local, regional or distant), new primary breast tumors (DCIS or invasive) or death due to any cause. For the current analysis, DFS was also the primary study endpoint, with follow-up starting at 2.5 years after randomization with a 10% margin. The secondary outcomes for this analysis were overall survival (OS), defined as time to death due to any cause starting at 2.5 years after randomization, and distant metastasis-free interval (DMFi), defined as time to distant recurrence starting at 2.5 years after randomization.

Statistical analysis

The analyses for primary and secondary outcomes (DFS, OS and DMFi) of the current study were performed using Kaplan Meier analysis, stratified for the type of endocrine therapy during the first 5 years, and nodal status at diagnosis. Hazard ratios (HRs) and treatment-by-marker interactions were estimated using Cox regression analysis.

Results

Cohorts

Of the 1824 postmenopausal patients enrolled in the IDEAL trial, 1339 were disease-free and on letrozole therapy at 2.5 years after randomization and were eligible for the current analysis. There were no significant differences in patient baseline characteristics between the randomized treatment arms in this subcohort (table 1).

Of the 438 patients who also participated in the TEAM trial, 311 patients were disease-free and on therapy at 2.5 years after randomization in the IDEAL study, and therefore eligible for our additional analysis. Patient characteristics of the IDEAL-only and IDEAL/TEAM patients are described in table 2. As compared to the IDEAL-only cohort (not participating in TEAM), IDEAL/TEAM patients were significantly older at randomization, more often treated with breast conserving therapy (55% vs 47.5%, X^2 p=0.037) and less often treated with chemotherapy (42.1 vs 77.6%) (table 2).

Regarding the prior endocrine therapy strategy, 816 IDEAL patients (60.9%) were treated with a sequential scheme of tamoxifen followed by an AI, 369 patients (27.6%) were treated with AI monotherapy, and 154 patients (11.5%) were treated with tamoxifen monotherapy. In the TEAM subgroup, 46.3% was treated with a sequential scheme, and 52.4% with AI monotherapy, as expected due to the TEAM trial design. Another four TEAM patients were treated with tamoxifen monotherapy because of refusal of switch to AI after 2.5 years of tamoxifen.

Main subgroup analysis in all patients

In the total selected IDEAL patient group (n=1339), 167 patients encountered a DFS event during follow-up (median follow-up of 7 years, including the first 2.5 years).

For node-negative patients, no benefit of longer endocrine therapy was found (HR 1.53, p=0.16, 95% CI 0.84-2.80). In contrast, for node-positive patients we observed a beneficial effect of longer extended therapy (HR 0.67, p=0.03, 95% CI 0.47-0.96), with a HR for interaction between nodal subgroups of 0.44 (95% CI 0.22-0.88, p=0.02) (table 3, figure 1).

Table 1: Characteristics of the IDEAL study cohort of patients who were disease-free and on therapy after 2.5 years of extended treatment.

N		2.5 years		5 years	
		%	N	%	
Age at randomisation	<55 years	191	28.6%	197	29.4%
	55-65 years	288	43.0%	283	42.2%
	65-75 years	151	22.6%	136	20.3%
	>75 years	39	5.8%	54	8.1%
Nodal status	pN0/pN0(i+)	176	26.3%	171	25.5%
	pN1(mi)/N1/N2/N3	493	73.7%	499	74.5%
Tumor type	ductal	508	75.9%	547	81.6%
	mucinous	5	.7%	6	.9%
	medullar	1	.1%	2	.3%
	lobular	113	16.9%	87	13.0%
	other/unknown	42	6.2%	28	4.2%
Histological grade	grade 1	115	17.2%	102	15.2%
	grade 2	278	41.6%	281	41.9%
	grade 3	205	30.6%	217	32.4%
	Gx	71	10.6%	70	10.4%
Progesteron receptor status	negative	113	16.9%	136	20.3%
	positive >=10%	528	78.9%	510	76.1%
HER2 status	Negative	242	36.2%	246	36.7%
	Positive	67	10.0%	63	9.4%
	unknown	360	53.8%	309	53.9%
Performed final surgery	breast conserving	335	50.1%	324	48.4%
	mastectomy	331	49.5%	344	51.3%
Prior chemotherapy	no	212	31.7%	198	29.6%
	yes	457	68.3%	472	70.4%
Prior endocrine treatment	5 years tamoxifen	76	11.4%	78	11.6%
	5 years AI	177	26.5%	192	28.7%
	2-3 years tam-> 3-2 years AI	416	62.2%	400	59.7%
Time after stop hormonal therapy (months)	0 to <6	602	90.0%	610	91.0%
	6 to <12	30	4.5%	27	4.0%
	12-27	37	5.5%	33	4.9%

When stratified for nodal status and type of endocrine therapy during the primary adjuvant therapy, we only observed the benefit of 5 years over 2.5 years of letrozole for node-positive patients in patients treated with prior sequential endocrine therapy (8 year DFS after randomization 89% vs 83.4%, HR 0.61, p=0.037, 95% CI 0.38-0.97) (figure 2). In this subgroup, the p-value for the treatment by subgroup interaction test based on nodal status was 0.05, indicating a significantly higher treatment effect in node-positive compared to node-negative patients. In all other considered subgroups, no benefit of longer extended therapy was observed (table 3).

Table 2: Characteristics of the IDEAL patients that participated earlier in the TEAM trial.

		Participation in TEAM trial				
		no		yes		X^2 p-value
		N	%	N	%	
Age at randomisation	<55 years	380	37.0%	8	2.6%	<0.001
	55-65 years	445	43.3%	126	40.5%	
	65-75 years	163	15.9%	124	39.9%	
	>75 years	40	3.9%	53	17.0%	
Nodal status	pN0/pN0(i+)	273	26.6%	74	23.8%	0.33
	pN1(mi)/N1/N2/N3	755	73.4%	237	76.2%	
Tumor type	ductal	803	78.1%	252	81.0%	0.84
	mucinous	9	.9%	2	.6%	
	medullar	2	.2%	1	.3%	
	lobular	160	15.6%	40	12.9%	
	other/unknown	54	5.3%	16	5.1%	
Histological grade	grade 1	161	15.7%	56	18.0%	0.06
	grade 2	422	41.1%	137	44.1%	
	grade 3	322	31.3%	100	32.2%	
	Gx	123	12.0%	18	5.8%	
Progesteron receptor status	negative	179	17.4%	70	22.5%	0.19
	positive >=10%	807	78.5%	231	74.3%	
HER2 status	negative	403	39.2%	85	27.3%	<0.001
	positive	125	12.2%	5	1.6%	
	unknown	500	48.6%	221	71.1%	
Performed final surgery	breast conserving	488	47.5%	171	55.0%	0.04
	mastectomy	535	52.0%	140	45.0%	
Prior chemotherapy	no	230	22.4%	180	57.9%	<0.001
	yes	798	77.6%	131	42.1%	
Prior endocrine treatment	5 years tamoxifen	150	14.6%	4	1.3%	<0.001
	5 years AI	206	20.0%	163	52.4%	
	2-3 years tam-> 3-2 years AI	672	65.4%	144	46.3%	
Time after stop hormonal therapy (months)	0 to <6	928	90.3%	284	91.3%	0.63
	6 to <12	43	4.2%	14	4.5%	
	12-27	57	5.5%	13	4.2%	

For the endpoint DMFi, similar results were observed (table 3). In node-positive patients previously treated with sequential therapy, a benefit of 5 years over 2.5 years of letrozole was shown (HR 0.50, p=0.03, 95% CI 0.27-0.94), but no differential effect between the treatment durations was observed for all other subgroups (p for interaction 0.14). For the endpoint OS, no benefit of longer extended therapy was shown for any of the subgroups (table 3).

Table 3: A subgroup analysis for the effect of 5 versus 2.5 years of extended letrozole on disease free survival (DFS), distant metastasis free interval (DMFi) and overall survival, stratified on prior endocrine therapy and nodal status.

		events	HR	p-value	95,0% CI			p for interaction		events	HR	p-value	95,0% CI			p for interaction
DFS	All patients								TEAM cohort							
All pre-treatments	No (n=347)	44	1.53	0.16	0.84	-	2.80	0.02	No (n=74)	16	2.12	0.15	0.77	-	5.85	0.06
	N+ (n=992)	123	0.67	0.03	0.47	-	0.96		N+ (n=237)	35	0.64	0.2	0.33	-	1.26	
5 years tamoxifen	No (n=48)	3	1.83	0.62	0.17	-	20.23	0.50	No (n=0)	-	-	-	-	-	-	-
	N+ (n=106)	11	0.87	0.81	0.26	-	2.85		N+ (n=4)	0	-	-	-	-		
5 years AI	No (n=102)	14	1.88	0.26	0.63	-	5.67	0.20	No (n=38)	9	1.31	0.70	0.34	-	5.08	0.84
	N+ (n=266)	35	0.81	0.52	0.41	-	1.57		N+ (n=125)	19	1.09	0.85	0.44	-	2.69	
2-3y tam-> 3-2y AI	No (n=196)	27	1.44	0.35	0.67	-	3.07	0.05	No (n=36)	7	3.68	0.12	0.71	-	18.97	0.02
	N+ (n=620)	77	0.60	0.03	0.38	-	0.95		N+ (n=108)	16	0.37	0.07	0.13	-	1.06	
DMFi																
All pre-treatments	No (n=347)	19	1.43	0.44	0.57	-	3.55	0.12	No (n=74)	5	1.89	0.49	0.31	-	11.40	0.19
	N+ (n=992)	67	0.63	0.06	0.38	-	1.03		N+ (n=237)	15	0.41	0.11	0.14	-	1.22	
5 years tamoxifen	No (n=48)	1	-	-	-		-	-	No (n=0)	0	-	-	-		-	-
	N+ (n=106)	6	1.10	0.91	0.22	-	5.45		N+ (n=4)	0	-	-	-		-	
5 years AI	No (n=102)	5	1.55	0.63	0.26	-	9.36	0.65	No (n=38)	2	1.18	0.91	0.07	-	19.95	0.98
	N+ (n=266)	16	0.98	0.96	0.37	-	2.60		N+ (n=125)	8	1.01	0.99	0.25	-	4.04	
2-3y tam-> 3-2y AI	No (n=196)	13	1.32	0.62	0.44	-	3.93	0.14	No (n=36)	3	2.56	0.44	0.23	-	28.30	0.09
	N+ (n=620)	45	0.50	0.03	0.27	-	0.94		N+ (n=108)	7	0.14	0.07	0.02	-	1.15	
OS																
All pre-treatments	No (n=347)	22	1.63	0.27	0.69	-	3.84	0.26	No (n=74)	8	2.07	0.32	0.49	-	8.71	0.32
	N+ (n=992)	70	0.89	0.61	0.55	-	1.42		N+ (n=237)	24	0.88	0.76	0.40	-	1.97	
5 years tamoxifen	No (n=48)	1	-	-	-		-	-	No (n=0)	0	-	-	-		-	-
	N+ (n=106)	8	1.80	0.42	0.43	-	7.55		N+ (n=4)	0	-	-	-		-	
5 years AI	No (n=102)	5	4.28	0.20	0.47	-	38.58	0.23	No (n=38)	3	2.13	0.55	0.18	-	24.88	0.65
	N+ (n=266)	20	0.96	0.92	0.40	-	2.30		N+ (n=125)	13	1.14	0.82	0.38	-	3.39	
2-3y tam-> 3-2y AI	No (n=196)	16	1.23	0.69	0.46	-	3.29	0.45	No (n=36)	5	2.11	0.41	0.35	-	12.66	0.34
	N+ (n=620)	42	0.75	0.36	0.41	-	1.39		N+ (n=108)	11	0.71	0.58	0.22	-	2.34	

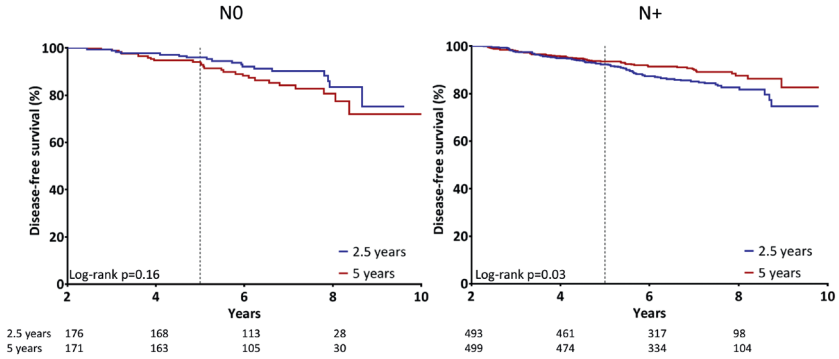


Figure 1: Kaplan-Meier analysis for disease-free survival of all patients that were disease-free and on therapy after 2.5 years, stratified for nodal status. Log-rank tests were used to assess the differences between treatment arms for each subgroup (reported as P values).

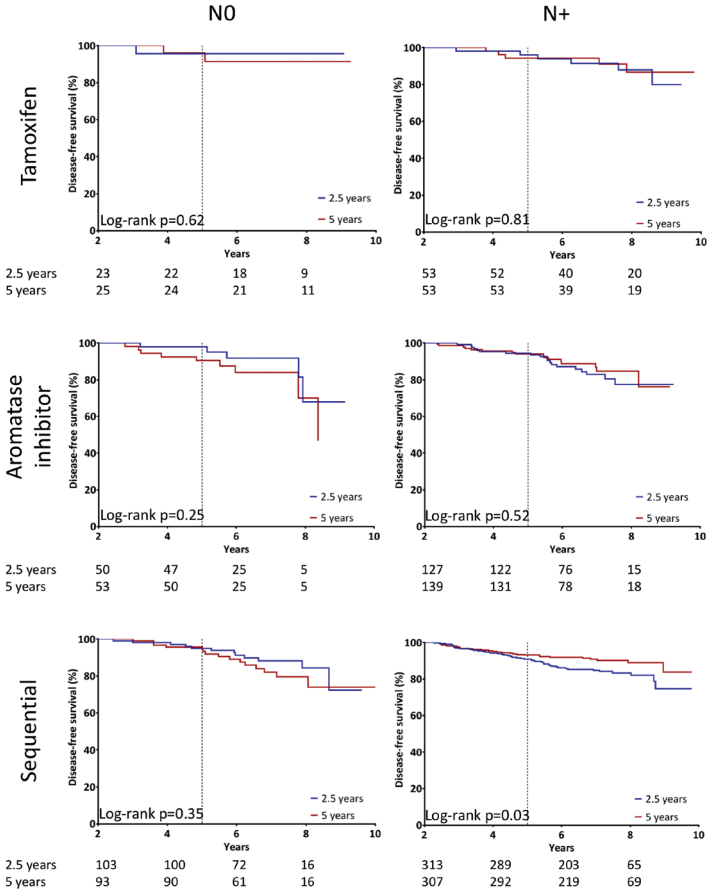


Figure 2: Kaplan-Meier analysis of the main analysis in all patients that were disease-free and on therapy after 2.5 years. Results are shown for disease-free survival, for the subgroups stratified on prior endocrine therapy and nodal status. Log-rank tests were used to assess the differences between treatment arms for each subgroup (reported as P values).

Additional subgroup analysis in TEAM patients

With respect to the additional analysis in the TEAM patient subgroup (n=311), 50 patients had a DFS event, of which 19 were DMFi events and 31 OS events. For DFS, a benefit of longer extended therapy was observed for node-positive patients pre-treated with sequential therapy, however without statistical significance (8 year DFS after randomization 90% vs 76.1%, HR 0.37, p=0.07, 95% CI 0.13-1.06). For DMFi, a similar non-significant benefit of longer therapy was found for the same subgroup (HR 0.14, p=0.07, 95% CI 0.02-1.15). Regarding OS, no benefit was shown for any of the subgroups (table 3).

Discussion

In this analysis of IDEAL patients, we found a significant benefit of longer (5 years, versus 2.5 years) extended letrozole therapy on disease-free and distant-metastasis free survival, for node-positive patients, and in particular those who received sequential adjuvant endocrine therapy during the first 5 years. In contrast, patients treated with AI monotherapy had no benefit of longer extended therapy, irrespective of nodal status. For overall survival no significant benefit of longer extended letrozole was observed in any subgroup, although the follow-up is relatively short for definite conclusions hereon.

The distribution of patients pre-treated with tamoxifen (followed by an AI) or with AI monotherapy in the full IDEAL cohort, might have been subject to allocation bias. Therefore, we performed an additional analysis in the IDEAL patients who also participated in the TEAM trial. Using the randomization of the TEAM trial, we balanced the previous endocrine therapy subgroups for baseline characteristics. In this additional analysis, similar numerical results were observed, although without statistical significance. This is most likely explained by the lack of power due to the smaller population size, and the low number of events in general. However, the similarity between the HRs for the total IDEAL cohort and the TEAM subgroup indicates that the results from the IDEAL cohort are not explained by an allocation bias.

The observation that (longer) extended therapy was only of value for node-positive patients, being at higher risk of recurrent disease, is in line with previous observations.

In a meta-analysis by Ibrahim et al, in which all patients were pre-treated with tamoxifen monotherapy, a subgroup analysis showed that the positive effect of extended endocrine therapy on breast cancer recurrence was only observed in node-positive patients (OR 0.70, 95% CI 0.58-0.84), and not in node-negative patients (OR 0.96, 95% CI 0.71-1.29).[6] Remarkably, in our analysis there was no benefit of longer extended therapy in either node-negative or node-positive patients that were treated with tamoxifen monotherapy. However, tamoxifen monotherapy for the first 5 years was not considered as standard therapy anymore during the conduct of the IDEAL trial, and most likely there might have been a selection bias of very-low risk patients who remain on tamoxifen after 2-3 years instead of switching to an AI. In these low-risk patients, a benefit of extended therapy is unlikely. Furthermore, tamoxifen monotherapy as prior endocrine therapy was a very small subgroup (12%) in the IDEAL trial, leading to a lack of power for conclusions in this subgroup.

The results of our analysis suggest that when patients were pre-treated with AI monotherapy for 5 years, there was no additional effect of 5 over 2.5 years of extended AI therapy. A possible explanation could be that the maximal treatment effect of aromatase inhibitors is reached after approximately 7.5 years. Therefore, after 5 years of AI monotherapy, an additional 5 years would have no benefit over 2.5 years. However, the results from this relative small subgroup analysis need to be interpreted with care, and should be validated in a meta-analytical setting before final conclusions can be drawn.

In all node-negative subgroups, there is a trend towards worse outcome for longer therapy, although none of these effects are statistically significant (table 3). For overall survival, this might be explained by the fact that in this low-risk subgroup, letrozole adverse events possibly leading to mortality outweigh the benefit of letrozole on breast cancer-related mortality. However, this does not explain why we see the same trend for longer therapy on distant metastasis-free interval. Further evaluation in larger analyses from collaborative groups, in the setting of a meta-analysis, are required to validate this effect.

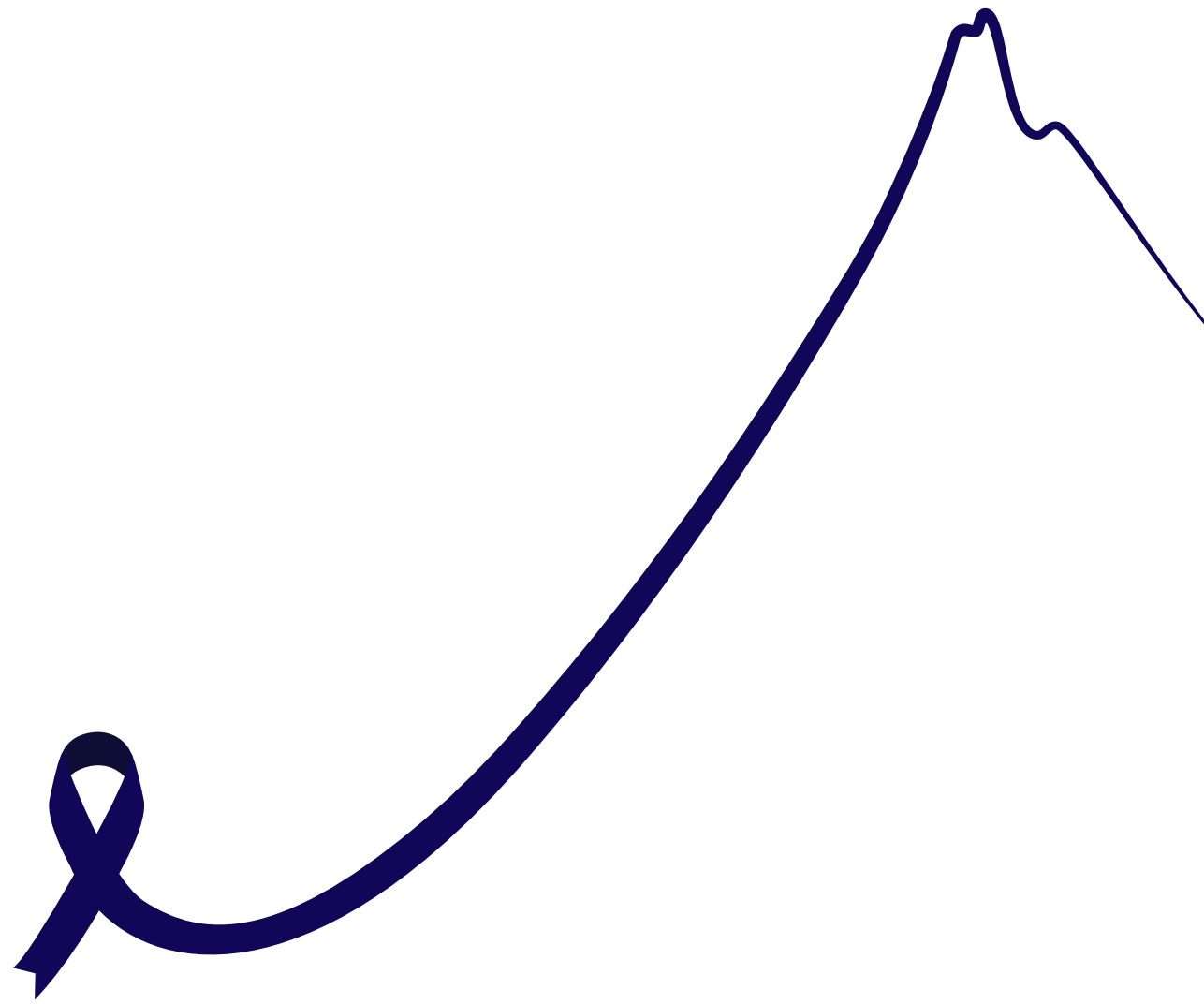
A limitation of this trial is that the analyses were performed in a subgroup of the original trial population, and this subgroup analysis was therefore not powered to detect small differences and might have suffered from multiple-testing error. Furthermore, In view of the design of the IDEAL study (having two treatment arms

and no placebo arm), it was not possible to investigate the value of extended therapy versus no extended therapy.

In conclusion, the results of the current exploratory analysis in IDEAL patients suggest that longer (versus shorter) extended endocrine therapy might be of value for node-positive patients, and in particular for those who were treated with tamoxifen followed by an AI for the first 5 years, which was not observed in the AI monotherapy subgroup. For all node-negative patients, there was no beneficial effect of longer therapy, and even a trend towards a worse outcome. Future studies, and future meta-analyses, are warranted to validate these results, and to further identify for which subgroup there is an effect of extended endocrine therapy after optimal endocrine therapy over the first 5 years.

References

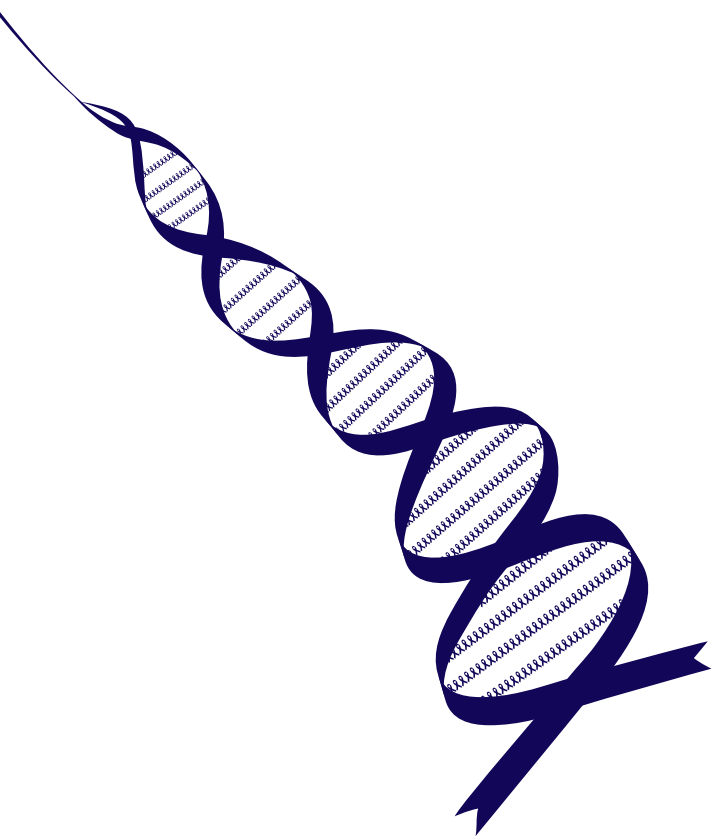
1. Derks MGM, Blok EJ, Seynaeve C et al. Adjuvant tamoxifen and exemestane in women with postmenopausal early breast cancer (TEAM): 10-year follow-up of a multicentre, open-label, randomised, phase 3 trial. The Lancet Oncology.
2. van de Velde CJ, Rea D, Seynaeve C et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-331.
3. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
4. Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).
5. Goss PE, Ingle JN, Martino S et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med 2003;349:1793-1802.
6. Ibrahim EM, Al-Hajeili MR, Bayer AM et al. Extended adjuvant endocrine therapy in early breast cancer: a meta-analysis of published randomized trials. Medical Oncology 2017;34(7):131.
7. Mamounas EP, Bandos H, Lembersky BC et al. Abstract S1-05: A randomized, double-blinded, placebo-controlled clinical trial of extended adjuvant endocrine therapy (tx) with letrozole (L) in postmenopausal women with hormone-receptor (+) breast cancer (BC) who have completed previous adjuvant tx with an aromatase inhibitor (AI): Results from NRG Oncology/NSABP B-42. Cancer Research 2017;77(4 Supplement):S1-05.
8. Tjan-Heijnen VC, Van Hellemond IE, Peer PG et al. Abstract S1-03: First results from the multicenter phase III DATA study comparing 3 versus 6 years of anastrozole after 2-3 years of tamoxifen in postmenopausal women with hormone receptor-positive early breast cancer. Cancer Research 2017;77(4 Supplement):S1-03.
9. Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg E et al. Optimal Duration of Extended Adjuvant Endocrine Therapy for Early Breast Cancer; Results of the IDEAL Trial (BOOG 2006-05). JNCI: Journal of the National Cancer Institute 2018;110(1):djx134.
10. Fontein DBY, Nortier JWR, Liefers GJ et al. High non-compliance in the use of letrozole after 2.5 years of extended adjuvant endocrine therapy. Results from the IDEAL randomized trial. European Journal of Surgical Oncology (EJSO) 2012;38(2):110-117.



Chapter 5

Treatment decisions and the impact of adverse events before and during extended endocrine therapy in postmenopausal early breast cancer

E.J. Blok
J.R. Kroep
W.M. Meershoek-Klein Kranenbarg
M. Duijm-de Carpentier
H. Putter
G.J. Liefers
J.W.R. Nortier
E.J.Th. Rutgers
C.S. Seynaeve
C.J.H. van de Velde



Accepted in:
European Journal of Cancer, 2018

Abstract

Background: Extended endocrine therapy beyond 5 years for postmenopausal breast cancer has been studied within multiple phase-III trials. Treatment compliance in these trials is generally poor. In this analysis, we aimed to determine which factors were associated with participation in the phase-III IDEAL trial, which factors are associated with early treatment discontinuation, and how this influenced survival outcome.

Methods: In the IDEAL trial, postmenopausal patients were randomized between 2.5 or 5 years of extended letrozole, after completing 5 years of endocrine therapy for hormone receptor-positive early breast cancer. A subgroup of this population participated earlier in the TEAM trial (5 years of exemestane, or 2.5 years of tamoxifen followed by exemestane as primary adjuvant therapy), in which we explored which factors were determinative for enrolment in the IDEAL study. In the IDEAL cohort, we evaluated which factors predicted for early treatment discontinuation, and the effect of early treatment discontinuation on disease-free survival (DFS).

Results: Nodal status, younger age and adjuvant chemotherapy were significantly associated with higher enrolment in the IDEAL trial. In the IDEAL cohort, adverse events, the type of primary endocrine therapy and the interval between primary and extended therapy were associated with early treatment discontinuation. Among the reported adverse events, depressive feelings (56%) was most frequently associated with early treatment discontinuation . Early treatment discontinuation was not associated with worse DFS (HR 1.02, 95%CI 0.76-1.37).

Conclusions: In this analysis, we found that risk factors were most strongly associated enrolment in the IDEAL trial. In contrast, patient experiences were the most significant factors leading to early treatment discontinuation, with no effect on DFS.

Introduction

Extended endocrine therapy for hormone receptor-positive (HR+) early breast cancer, beyond the standard 5 years, is more frequently being used over the recent years. After five years of tamoxifen, it has been shown that extended therapy with either 5 additional years of tamoxifen or 5 years of an aromatase inhibitor (AI) has clinical benefit, in particular in high-risk (lymph node positive) disease. [1-5] However, less data are available on the value of extended adjuvant endocrine therapy after primary adjuvant therapy including an AI. Recently, results of trials studying extended therapy after an optimal primary adjuvant regimen (including AIs during the first 5 years) were presented and partly published, showing no significant benefit of extended therapy for the total group, and suggesting that mainly high-risk subgroups might benefit from this extended therapy. [6, 7]

One of the problems regarding adjuvant endocrine therapy, both during primary and extended therapy, is early treatment discontinuation. In the trials reporting on 5 years of extended therapy, patient compliance (finishing 5 years of extended therapy) was consistently low. [4, 6-8] In most studies this is considered to be attributed to the side effects of endocrine therapy. However, in the placebo-controlled NSABP B42 trial, in which patients after 5 years of AI-based therapy were randomized between 5 years of letrozole or placebo, the early treatment discontinuation in the placebo- arm was similar as in the letrozole-group (62% and 60% on therapy at 5 years). [8] A similar effect was observed in the MA.17 trial, in which patients were randomized between 5 years of letrozole or placebo after 5 years of tamoxifen, reporting a discontinuation rate of 9.9% vs 9.8% in the letrozole and placebo groups respectively. [3, 9] This indicates that other factors than actual treatment toxicity also play a role in the early discontinuation of extended adjuvant endocrine therapy. Knowledge of these factors would enable the clinician and health care workers to tailor the support of patients during extended adjuvant endocrine therapy, in order to decrease early treatment discontinuation.

In the Dutch 'Investigation on the Duration of Extended Adjuvant Letrozole' (IDEAL) trial in which postmenopausal patients were randomly allocated to either 2.5 or 5 years of letrozole after 5 years of any type of adjuvant endocrine therapy, 629 (35%) patients stopped therapy earlier than planned (27% in 2.5 years group, 43% in the five-year group), of which the majority (59%) reported side effects as the main reason for

early treatment discontinuation. [7] A preliminary evaluation after 2.5 years of follow-up suggested that nodal status, type of earlier endocrine therapy and the interval between primary and extended adjuvant therapy could influence patient compliance [10] A number of patients in the IDEAL trial earlier participated in the Tamoxifen Exemestane Adjuvant Multinational (TEAM) trial, in which postmenopausal early breast cancer patients were randomized between either 5 years of adjuvant exemestane, or a sequential scheme of tamoxifen for 2.5 years followed by 2.5 years of exemestane. [11, 12] At that time, extended adjuvant therapy was not yet standard of care in the Netherlands, but Dutch TEAM patients were allowed to be enrolled in the IDEAL study thereafter.

The present analyses were performed to explore factors contributing to treatment decisions at enrolment in, and during the course of the IDEAL trial. The first aim was to identify which factors were associated with enrolment in the IDEAL trial after participation in the TEAM trial. The second aim was to identify which baseline factors were associated with early treatment discontinuation in the IDEAL trial. The third aim was to assess which specific adverse events are associated with either the decision to extend endocrine therapy after TEAM participation, and early treatment discontinuation during the IDEAL trial. Finally, we investigated the effect of adverse event-based early treatment discontinuation on disease-free survival (DFS).

Methods

Patient cohorts and data collection

In the IDEAL trial, 1824 postmenopausal early HR+ breast cancer patients were randomly allocated to either 2.5 or 5 years of letrozole, after receiving any type of 5-years primary adjuvant endocrine therapy. The eligibility criteria were: completion of 5 years (±3 months) of primary adjuvant endocrine therapy (either within a regular or clinical trial setting), end of primary adjuvant therapy within 24 months before inclusion, and disease-free status without other malignancies at the time of randomisation.

In the international TEAM trial, 2754 Dutch patients were included of whom 2363 were treated in a center that later on also participated in the IDEAL trial. The flowchart of the selection procedure used for determination of the cohort of TEAM

patients being theoretically eligible for the IDEAL study and the current analyses is depicted in figure 1. The selection criteria included: be treated in a center that later on would also participate in IDEAL, be free of any recurrence or new primary breast tumour at the end of active TEAM trial medication, completion of 5 years (± 3 months) of endocrine therapy within the TEAM trial, no other malignancy within previous 5 years, and available follow-up data until the time of randomization in the IDEAL (not lost to follow-up). Using these criteria, we defined a cohort of 1216 patients that were theoretically eligible for IDEAL enrolment after TEAM participation (figure 1).

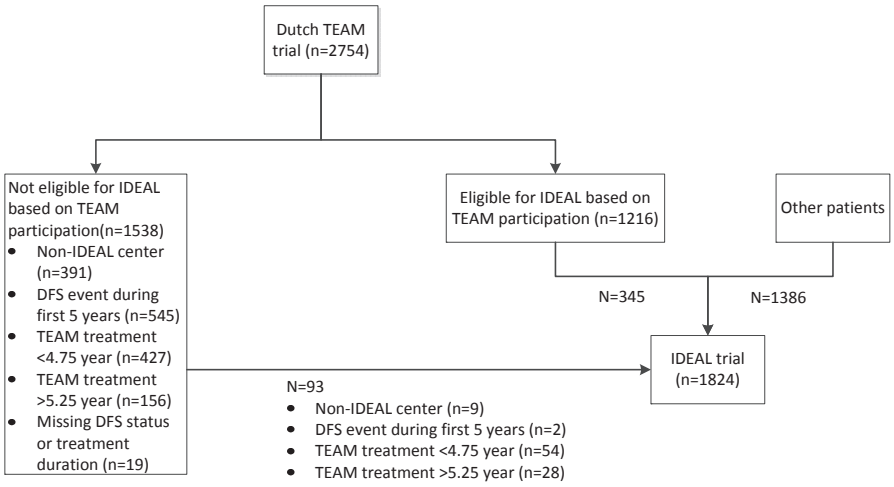


Figure 1 – An overview of the relation between the TEAM and IDEAL trial patient cohorts

For both studies, data were locally registered on the clinical record forms (CRFs) and centrally collected at the LUMC Datacenter Department of Surgery. Details on data collection for both studies were reported earlier [7, 12] Relevant for the current analysis is the datacollection at baseline of patient, tumor and treatment characteristics, the registration of adverse events (AEs) during active treatment using the Common Terminology Criteria for Adverse Events (TEAM: version 2.0, IDEAL: version 4.0), and registration of details on the stop of therapy (date and reason, "stopped due to adverse event" was an option to select on the CRF).

Data handling and endpoints

All adverse events, irrespective of grade, were considered for the analyses. Cardiac disorders were combined into categories based on the CTC categories ('cardiovascular general' and 'cardiovascular arrhythmia' in CTC version 2.0, and 'cardiac disorders' in CTC version 4.0). For the analyses on AEs, the most frequently occurring AEs and most relevant for the given endocrine therapy were included, as based on the primary publications of both trials. [7,12]

Early treatment discontinuation was defined as a definite stop of trial-based letrozole before the designated treatment stop date as reported by the treating physician on the CRF, except when the medication stop was caused by recurrence or death. Some patients stopped letrozole in the IDEAL study, and continued with a different aromatase inhibitor off-trial. These patients were still regarded as being early stopped with therapy for all analyses, since registration of off-trial medication was not assured for all patients.

The primary endpoint of the IDEAL trial was disease-free survival (DFS), defined as the time from randomization to recurrence (either local, regional or distant), new primary breast tumors (DCIS or invasive) or death due to any cause. [7]

Statistical analysis

In the cohort of TEAM patients that was eligible for inclusion in the IDEAL trial, baseline clinicopathological parameters were tested for association with enrolment in the IDEAL. Statistical independence was tested using Chi-square (X^2) tests with an alpha of 0.05. Furthermore, we determined the percentage of patients that continued with extended therapy in the IDEAL, for the patients who reported the most frequent and relevant adverse events associated with endocrine therapy.

In the cohort of IDEAL patients, baseline clinicopathological parameters were tested for association with the occurrence of adverse events and with early treatment discontinuation, using Chi-square (X^2) tests with an alpha of 0.05 and multivariate logistic regression modelling. Furthermore, for each selected (most frequent) adverse event we determined the percentage of patients that stopped therapy because of adverse events, thereby showing which adverse events are related to early treatment discontinuation . In the patients who encountered and adverse event, Kaplan-Meier and univariate stratified Cox-regression analysis(stratified for earlier endocrine therapy, chemotherapy, time between regular and extended therapy, and nodal

status) were used to determine the effect of early treatment discontinuation due to an adverse event on DFS. A Cox-regression interaction test was used to test for interaction between treatment arms.

Statistical analysis was performed using SPSS 23.0 (IBM), and a 2-sided p-value of 0.05 or lower was regarded as statistically significant for each analysis.

Results

Of the Dutch TEAM cohort, 1216 patients were theoretically eligible for participation in the IDEAL study (figure 1), of which 345 indeed were enrolled in IDEAL (28%). As described in table 1, we observed that patients who were younger, received chemotherapy, underwent an axillary dissection or had node-positive disease, were significantly more likely to be enrolled in the IDEAL study and therefore to continue with extended endocrine therapy (table 1). The occurrence of an adverse event (AE) during the TEAM trial was not associated to participation in the IDEAL trial (table 1).

Table 1 – Baseline clinicopathological characteristics of TEAM patients eligible for participation in the IDEAL trial. Multivariate logistic regression modelling used to determine odds ratios.

		IDEAL study participation						
		no		yes				
		N	%	N	%	X^2 P-value	Odds ratio	P-value
Randomisation	5y exemestane	489	71.2%	198	28.8%	0.69	1.00	
	2.5y tamoxifen, 2.5y exemestane	382	72.2%	147	27.8%		0.83	0.09
Age	<50 years	12	46.2%	14	53.8%	<0.001	1.00	
	50-59	270	67.7%	129	32.3%		0.61	0.13
	60-69	309	69.3%	137	30.7%		0.57	0.10
	>=70	280	81.2%	65	18.8%		0.31	0.00
ECOG Performance Status**	0	682	70.9%	280	29.1%	0.22	1.00	
	1	150	77.7%	43	22.3%		0.59	0.003
	unknown	16	66.7%	8	33.0%		1.43	0.12
Tumour type	ductal	645	71.5%	257	28.5%	0.45	1.00	
	lobular	128	69.2%	57	30.8%		1.71	0.50
	mixed ductal/ lobular	43	74.1%	15	25.9%		1.90	0.42
	other	41	82.0%	9	18.0%		1.94	0.42
	NOS adenocarci-noma	8	80.0%	2	20.0%		1.59	0.62

Table 1 continued

		IDEAL study participation						
		no		yes				
		N	%	N	%	X^2 P-value	Odds ratio	P-value
Bloom Richardson grade	Grade 1	139	68.5%	64	31.5%	0.54	1.00	
	Grade 2	408	72.6%	154	27.4%		0.80	0.17
	Grade 3	260	71.6%	103	28.4%		0.77	0.15
Tumoursize	T1	427	71.6%	169	28.4%	0.24	1.00	
	T2	400	71.7%	158	28.3%		0.86	0.21
	T3	27	64.3%	15	35.7%		1.02	0.96
	T4	15	88.2%	2	11.8%		0.22	0.04
Nodal status	N-	278	79.2%	73	20.8%	<0.001	1.00	
	N+	593	68.6%	272	31.4%		1.71	0.008
Hormonereceptor Status ER/PgR	ER+/PgR+	637	71.7%	252	28.3%	0.06	1.00	
	ER+/PgR-	154	67.2%	75	32.8%		0.98	0.87
	ER+/PgR unknown	65	80.2%	16	19.8%		0.64	0.11
	ER-/PgR+	13	92.9%	1	7.1%		0.23	0.05
	ER-/PgR-*	2	100.0%	0	0.0%		0.00	1.00
HER2 status	negative	730	72.9%	272	27.1%	0.28	1.00	
	positive	86	68.3%	40	31.7%		1.21	0.29
Adjuvant Chemotherapy	no	649	76.0%	205	24.0%	<0.001	1.00	
	yes	222	61.3%	140	38.7%		1.25	0.14
Most Extensive Surgery	mastectomy	450	72.6%	170	27.4%	0.61	1.00	
	wide local excision	420	70.6%	175	29.4%		1.39	0.008
Axillary Dissection	no	211	76.7%	64	23.3%	0.03	1.00	
	yes	660	70.1%	281	29.9%		0.97	0.90
Adverse event during treatment	no	264	71.5%	105	28.5%	0.97	1.00	
	yes	607	71.7%	240	28.3%		1.03	0.84

*Ineligible for inclusion

NOS=Not otherwise specified

**0 = asymptomatic

1 = symptomatic, full ambulatory

Regarding inclusion in the IDEAL study in relation to AEs experienced during the TEAM study medication, we considered hot flashes, fatigue, arthralgia, lymphedema, nausea and alopecia (most frequently occurring), together with relevantly endocrine therapy-associated adverse events (depressive feelings, fractures and cardiac disease) (table 2). Of these factors, only cardiac arrhythmias showed a lower participation rate in the IDEAL trial (table 2).

Table 2 – Most frequently reported adverse events in the TEAM trial, associated with participation in the extended endocrine therapy IDEAL trial.

	IDEAL study participation			
	no		yes	
	N	%	N	%
Hot flashes	277	68,6%	127	31,4%
Fatigue	208	78,5%	57	21,5%
Arthralgia	110	69,2%	49	30,8%
Cardiovascular disease (general)	125	71,8%	49	28,2%
Lymphedema	116	75,8%	37	24,2%
Nausea	57	71,3%	23	28,8%
Alopecia	45	70,3%	19	29,7%
Fracture	51	78,5%	14	21,5%
Cardiovascular disease (arrhythmia)	45	83,3%	9	16,7%
Mood alteration-depression	32	71,1%	13	28,9%

Regarding the question which factors contributed to early letrozole discontinuation, we observed that patients who received tamoxifen monotherapy as primary adjuvant therapy, or had a longer break between initial and extended therapy reported significantly more frequently adverse events, and also stopped therapy significantly more often due to adverse events or treatment refusal (table 3). Of all patients reporting one or more adverse events, 31% stopped therapy early, compared to 6.6% of patients that did not report an adverse event (table 3).

We explored the most frequently reported and most relevantly AI-associated adverse events [7] , and determined the percentage of patients that stopped letrozole therapy early due to adverse events. The most pronounced specific AEs associated with early treatment discontinuation were depressive feelings (55.8% of patients reporting a depression stopped due to an adverse event), alopecia (47.4%) and arthralgia (44.3%). In contrast, osteoporosis (14.1%), fractures (12.9%) and back pain (11.9%) were seldom associated with early treatment continuation due to adverse events (figure 2).

For patients who stopped letrozole therapy early due to adverse events, the mean treatment duration was 13 months (8 months for patients randomized to 2.5 years, and 18 months for patients randomized to 5 years). In contrast, in patients who did not stop due to an adverse event, the average treatment duration was 42 months (30 months for 2.5 years group, and 55 months for 5 years group). The effect of this shorter therapy duration on DFS was studied comparing patients who stopped early due to an AE with patients who did not stop early due to an AE (figure 3). Early treatment

Table 3 – Clinicopathological baseline factors in the IDEAL trial patients, associated with both the reporting of adverse events, and to early treatment discontinuation. Multivariate logistic regression modelling used to determine odds ratios.

		Reported adverse event					Early treatment discontinuation						
		no		yes			no		yes				
		N	%	N	%	X^2 P-value	N	%	N	%	X^2 P-value	Odds ratio	P-value
Allocated treatment	2.5 years	262	29.0%	640	71.0%	0.69	723	80.2%	179	19.8%	<0.001	1.00	
	5 years	255	28.2%	649	71.8%		649	71.8%	255	28.2%		1.71	**<0.001**
Age at randomisation	<55 years	144	28.4%	363	71.6%	0.42	385	75.9%	122	24.1%	0.28	1.00	
	55-65 years	226	29.9%	531	70.1%		585	77.3%	172	22.7%		0.95	0.74
	65-75 years	104	25.7%	301	74.3%		294	72.6%	111	27.4%		1.08	0.73
	>75 years	43	31.4%	94	68.6%		108	78.8%	29	21.2%		0.77	0.41
Tumour type	ductal	424	30.2%	982	69.8%	0.04	1071	76.2%	335	23.8%	0.60	1.00	
	mucinous	3	18.8%	13	81.3%		10	62.5%	6	37.5%		3.78	**0.04**
	medullar	2	28.6%	5	71.4%		5	71.4%	2	28.6%		1.97	0.47
	lobular	62	21.3%	229	78.7%		217	74.6%	74	25.4%		1.13	0.52
	other	26	30.6%	59	69.4%		68	80.0%	17	20.0%		0.72	0.32
Histological grade	grade 1	87	30.5%	198	69.5%	0.69	217	76.1%	68	23.9%	0.39	1.00	
	grade 2	218	28.4%	550	71.6%		575	74.9%	193	25.1%		1.08	0.68
	grade 3	174	31.0%	387	69.0%		436	77.7%	125	22.3%		0.91	0.64
T-stage	T1	245	29.7%	580	70.3%	0.28	613	74.3%	212	25.7%	0.25	1.00	
	T2	227	27.3%	603	72.7%		645	77.7%	185	22.3%		0.85	0.23
	T3	23	24.5%	71	75.5%		68	72.3%	26	27.7%		1.10	0.75
	T4	15	38.5%	24	61.5%		32	82.1%	7	17.9%		0.66	0.44
N-stage	N0	135	29.0%	331	71.0%	0.98	349	74.9%	117	25.1%	0.07	1.00	
	N1	307	28.7%	761	71.3%		801	75.0%	267	25.0%		1.00	0.99
	N2	55	27.6%	144	72.4%		163	81.9%	36	18.1%		0.69	0.19
	N3	16	27.6%	42	72.4%		49	84.5%	9	15.5%		0.57	0.21

Progesteron receptor status	not done	16	24.2%	50	75.8%	0.55	52	78.8%	14	21.2%	0.73	1.00	
	negative	103	30.4%	236	69.6%		261	77.0%	78	23.0%		1.05	0.89
	positive >=10%	396	28.3%	1002	71.7%		1056	75.5%	342	24.5%		1.13	0.73
Prior endocrine treatment	5 years tamoxifen	50	22.7%	170	77.3%	0.006	153	69.5%	67	30.5%	0.015	1.00	
	5 years AI	133	25.5%	389	74.5%		388	74.3%	134	25.7%		0.71	0.11
	2-3 years tam-> 3-2 years AI	334	31.4%	730	68.6%		831	78.1%	233	21.9%		0.64	**0.02**
Time after stop hormonal therapy	0 to <6 months	473	29.5%	1132	70.5%	0.04	1242	77.4%	363	22.6%	<0.001	1.00	
	6 to <12 months	24	26.1%	68	73.9%		62	67.4%	30	32.6%		1.71	**0.04**
	12 to 24 months	20	18.3%	89	81.7%		68	62.4%	41	37.6%		1.91	**0.01**
Prior chemotherapy	no	150	26.2%	423	73.8%	0.12	420	73.3%	153	26.7%	0.07	1.00	
	yes	367	29.8%	866	70.2%		952	77.2%	281	22.8%		0.87	0.42
Performed final surgery baseline	breast conserving	273	31.0%	609	69.0%	0.03	667	75.6%	215	24.4%	0.79	1.00	
	mastectomy	242	26.30%	677	73.70%		700	76.2%	219	23.8%		0.97	0.83
Axillary dissection baseline	not done	106	28.90%	261	71.10%	0.90	279	76.0%	88	24.0%	0.98	1.00	
	done	411	28.60%	1028	71.40%		1093	76.0%	346	24.0%		1.04	0.86
Reported adverse event	no	-	-	-	-	-	483	93.4%	34	6.6%	<0.001	1.00	
	yes	-	-	-	-	-	889	69.0%	400	31.0%		5.50	**<0.001**

discontinuation due to an AE had no effect on DFS (HR 1.02, 95%CI 0.76-1.37, p=0.90). This effect was similar in both IDEAL therapy arms (HR for interaction 0.85, 95%CI 0.47-1.51, p=0.57).

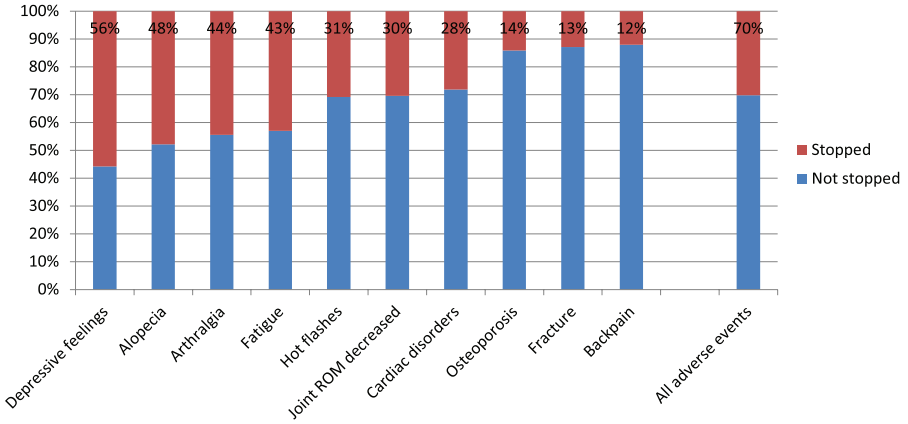


Figure 2 – Most frequently reported adverse events in the IDEAL trial, associated with early treatment discontinuation. For each adverse event, the percentage of patients that stopped therapy early due to adverse events is shown.

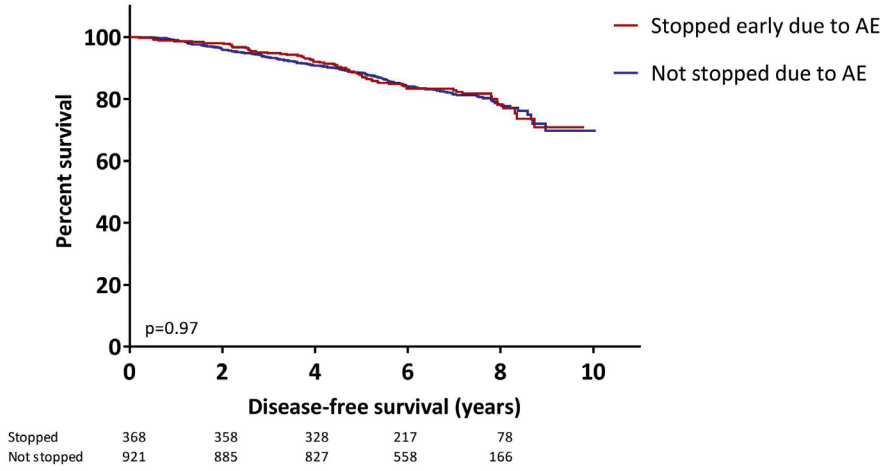


Figure 3 –Kaplan Meier survival analysis in the group of patients who encountered an adverse event, comparing patients who stopped therapy early due to an AE with patients who did not stop early due to an AE. Log-rank test is used to determine the p-value.

Discussion

In this explorative analysis, we found that classical risk factors for recurrent disease such as nodal status were more associated with starting extended endocrine therapy in the IDEAL study, whereas patient-centred factors like adverse events and experiences with earlier endocrine therapy were more driving factors regarding early treatment discontinuation.

We observed that the factors associated with enrolment in the IDEAL trial after completion of the TEAM study medication, were in general risk factors for disease recurrence (like nodal status and previous adjuvant chemotherapy) as well as a younger age. Remarkably, the occurrence of adverse events during primary adjuvant treatment (not leading to early stopping of endocrine therapy) did not appear to influence enrolment in the IDEAL trial. Only cardiac events during the TEAM trial, were associated with a lower rate of enrolment in the IDEAL trial, which can be explained by the fact that uncontrolled cardiovascular disease was an exclusion criterion for the IDEAL trial. The decision whether or not to extend endocrine therapy, was most likely a risk/benefit-based decision. In contrast, exploration of factors associated with early discontinuation of extended endocrine therapy revealed that the occurrence of adverse events played a major role in early treatment discontinuation. Without adverse events, 6.6% of patients stopped early, whereas 31% with an adverse event stopped prematurely. Other factors with impact on early treatment discontinuation were the type of primary adjuvant endocrine therapy, the interval between primary and extended therapy, and the allocated treatment. The latter could be explained by the fact that patients who were allocated to 5 years of extended letrozole, had twice as much time to stop therapy early, compared to the patients who were allocated to 2.5 years of letrozole.

In general, there is concordance between the reporting of adverse events in a particular subgroup, and the rate of early treatment discontinuation. For example, tamoxifen monotherapy and a longer interval between regular and extended therapy, are both associated with the reporting of adverse events and with early treatment discontinuation. For both of these factors, the explanation for a higher rate of reported adverse events might be that the patients were not familiar (anymore) with the side effects of aromatase inhibitors, and therefore reported them more often and were more likely to stop therapy due to these complaints. The only discordance in the association

between adverse events and early treatment discontinuation is the type of surgery. Patients who underwent a mastectomy reported significantly more adverse events compared to patients treated with a lumpectomy, but did not differ with regard to early treatment discontinuation rate. Most likely, adverse events in the mastectomy group were not associated with endocrine therapy but with the local therapy itself, and therefore did not lead to a difference in early endocrine treatment discontinuation.

Of the adverse events associated with early treatment discontinuation, arthralgia, fatigue, alopecia and in particular depressive feelings/mood alterations were the most pronounced. These adverse events probably caused the highest (subjective) impact on the quality of life. Remarkably, osteoporosis and fractures were associated with a lower rate of early treatment discontinuation, while these are well known side effects from AI therapy. Apparently, osteoporosis has less impact on the life of patients, and is not considered to be a major problem as it can be treated with calcium, vitamin D and bisphosphonates.

In the patients who reported one or more adverse events, early treatment discontinuation did not lead to a worse disease-free survival. On one hand this is remarkable, since it has been suggested that patients who stop earlier than planned, are at a higher risk for recurrent disease. [13, 14] However, these latter studies were performed in the context of standard (5 years) adjuvant endocrine therapy, in which it has been shown that 5 years of therapy is superior over 2 years of therapy. [15] On the other hand, our data are derived from the IDEAL patient cohort, and the primary evaluation of this trial did not show a difference in DFS between 2.5 and 5 years of extended letrozole either. [7] Our results therefore support that for extended therapy with letrozole after optimal primary adjuvant endocrine therapy, a longer extended treatment does not lead to an improved survival for the total group. However, in view of this unplanned analysis, it is warranted to study this further in collaboration with other extended endocrine therapy trials.

We observed that a number of patients (n=93) were enrolled in both the TEAM and IDEAL trial, without being part of this IDEAL-eligible TEAM-population. However, this doesn't mean that they were ineligible for the IDEAL. For example, patients could have went off-study in the TEAM trial, but continued with endocrine therapy outside the TEAM-trial, becoming then eligible for the IDEAL study after 5 years of endocrine therapy.

A limitation of this analysis is that when a patient experienced multiple adverse events, and stopped due to the occurrence of an adverse event, it was not registered which specific adverse event was responsible for the discontinuation of therapy. Another limitation is that adverse events were only collected during active treatment in the IDEAL trial. Therefore, the registration period for the 5-years arm was twice as long as the 2.5 years arm, and quantitative comparisons regarding specific adverse events cannot be made.

In summary, our study demonstrates that the decision to enrol patients in an extended endocrine therapy trial is mainly influenced by high-risk factors and a sufficiently expected benefit, reflecting a risk-/benefit-based decision making process. In contrast, compliance to complete extended endocrine therapy is mainly influenced by adverse events, and type and timing of earlier therapy. Noteworthy, early treatment discontinuation due to adverse events did not lead to a worse DFS in our analysis, which should be investigated further. In our opinion, these results can contribute towards a better understanding of and improving early treatment discontinuation of extended endocrine therapy both in clinical trials and in standard care.

Registration

This trial is registered in the Netherlands with the Netherlands Trial Register, NTR3077, the Dutch Breast Cancer Research Group (BOOG 2006-05) and Eudra-CT 2006-003958-16.

Funding

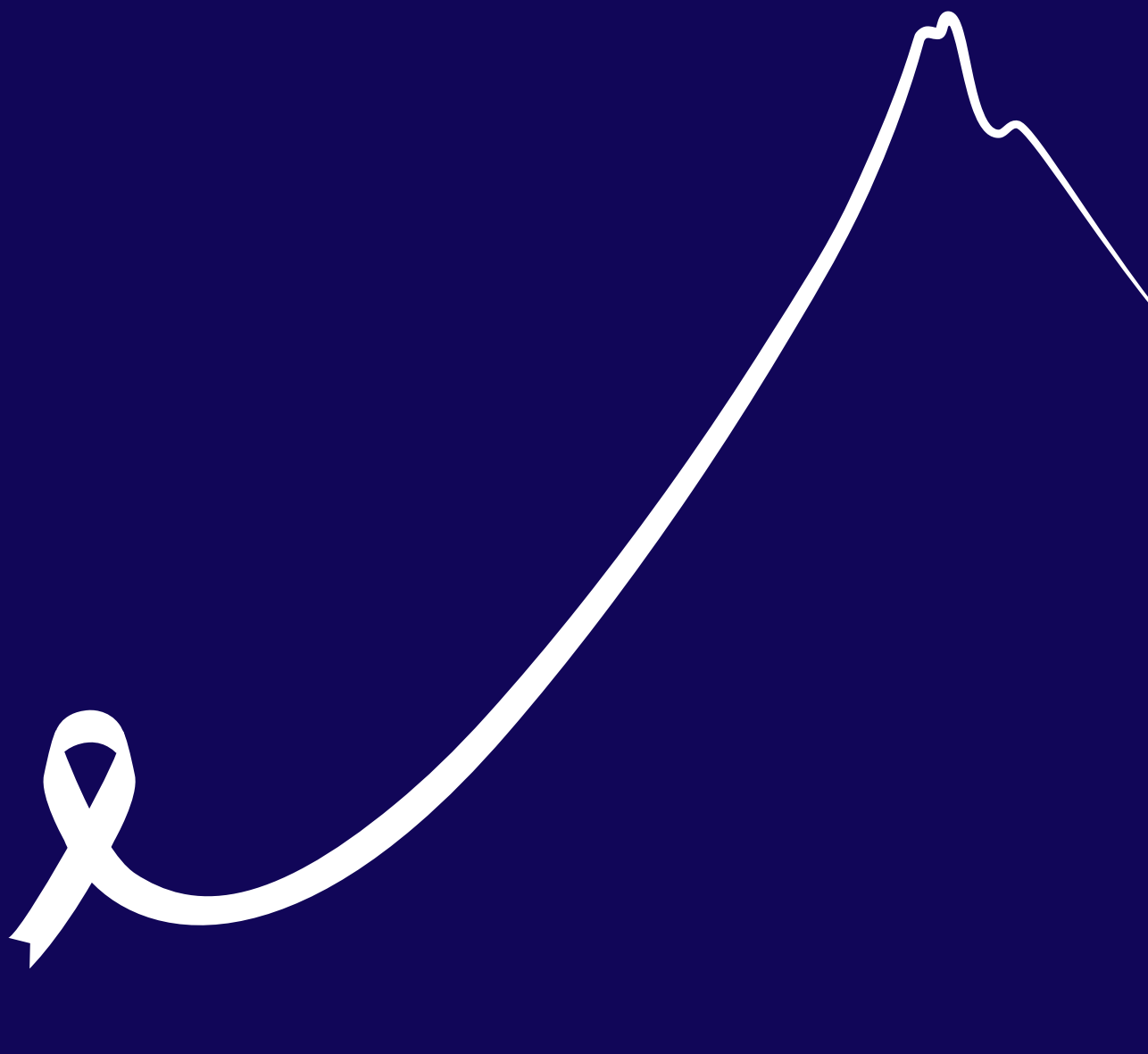
The IDEAL trial was supported by Novartis by means of an independent educational grant (CFEM345DNL03), although this specific analysis was not financially supported.

Conflicts of interest

The funding body was not involved in collection, analysis, or interpretation of the data, nor in the decision to submit for publication. The corresponding author confirms that he had access to all data and had final responsibility for the decision to submit for publication. All authors report no financial or personal disclosures related to this work.

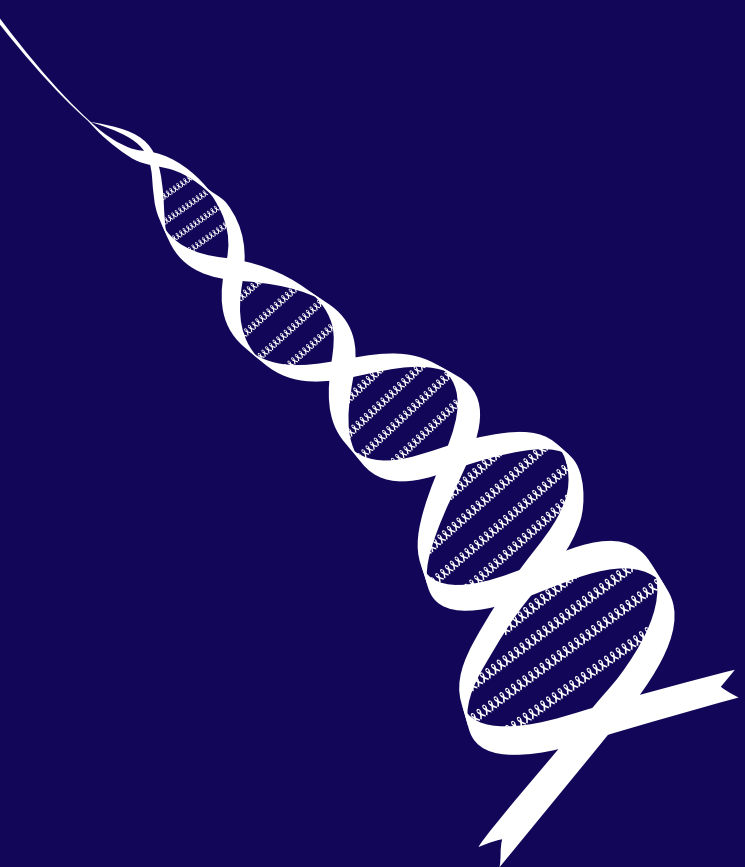
References

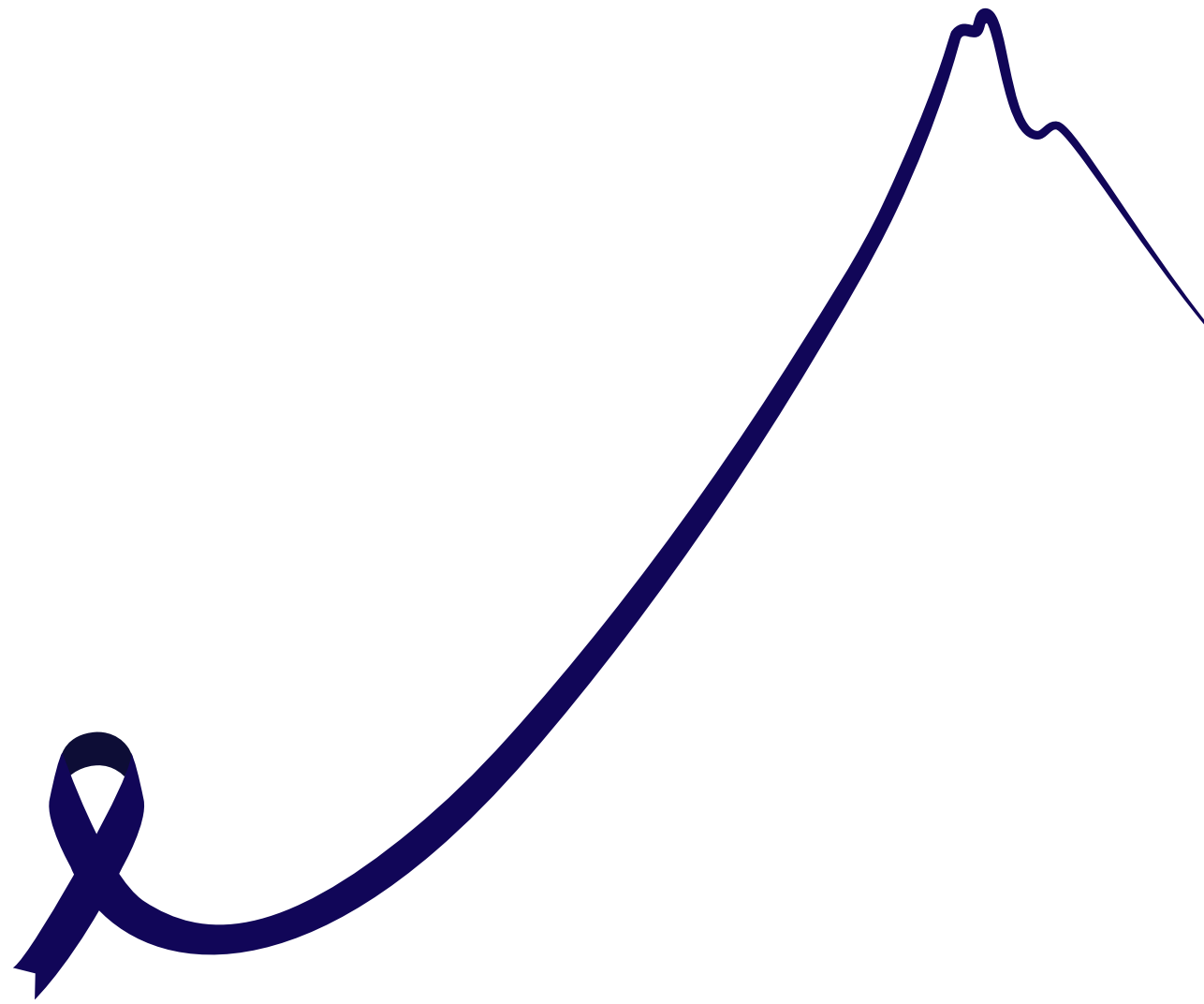
1. Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.
2. Ibrahim EM, Al-Hajeili MR, Bayer AM et al. Extended adjuvant endocrine therapy in early breast cancer: a meta-analysis of published randomized trials. Medical Oncology 2017;34(7):131.
3. Goss PE, Ingle JN, Martino S et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med 2003;349:1793-1802.
4. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381:805-816.
5. Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).
6. Tjan-Heijnen VCG, van Hellemond IEG, Peer PGM et al. Extended adjuvant aromatase inhibition after sequential endocrine therapy (DATA): a randomised, phase 3 trial. The Lancet Oncology 2017;18(11):1502-1511.
7. Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg E et al. Optimal Duration of Extended Adjuvant Endocrine Therapy for Early Breast Cancer; Results of the IDEAL Trial (BOOG 2006-05). JNCI: Journal of the National Cancer Institute 2018;110(1):djx134.
8. Mamounas EP, Bandos H, Lembersky BC et al. Abstract S1-05: A randomized, double-blinded, placebo-controlled clinical trial of extended adjuvant endocrine therapy (tx) with letrozole (L) in postmenopausal women with hormone-receptor (+) breast cancer (BC) who have completed previous adjuvant tx with an aromatase inhibitor (AI): Results from NRG Oncology/NSABP B-42. Cancer Research 2017;77(4 Supplement):S1-05.
9. Hadji P. Improving compliance and persistence to adjuvant tamoxifen and aromatase inhibitor therapy. Critical Reviews in Oncology/Hematology 2010;73(2):156-166.
10. Fontein DBY, Nortier JWR, Liefers GJ et al. High non-compliance in the use of letrozole after 2.5 years of extended adjuvant endocrine therapy. Results from the IDEAL randomized trial. European Journal of Surgical Oncology (EJSO) 2012;38(2):110-117.
11. Derks MGM, Blok EJ, Seynaeve C et al. Adjuvant tamoxifen and exemestane in women with postmenopausal early breast cancer (TEAM): 10-year follow-up of a multicentre, open-label, randomised, phase 3 trial. The Lancet Oncology 2017;18(9):1211-1220.
12. van de Velde CJ, Rea D, Seynaeve C et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-331.
13. Ulcickas Yood M, Owusu C, Buist DSM et al. Mortality Impact of Less-than-Standard Therapy in Older Breast Cancer Patients. Journal of the American College of Surgeons 2008;206(1):66-75.
14. Hsieh KP, Chen LC, Cheung KL et al. Interruption and Non-Adherence to Long-Term Adjuvant Hormone Therapy Is Associated with Adverse Survival Outcome of Breast Cancer Women - An Asian Population-Based Study. PLoS ONE 2014;9(2):e87027.
15. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-1717.



Part 2

Biomarker-based tailoring of adjuvant therapy



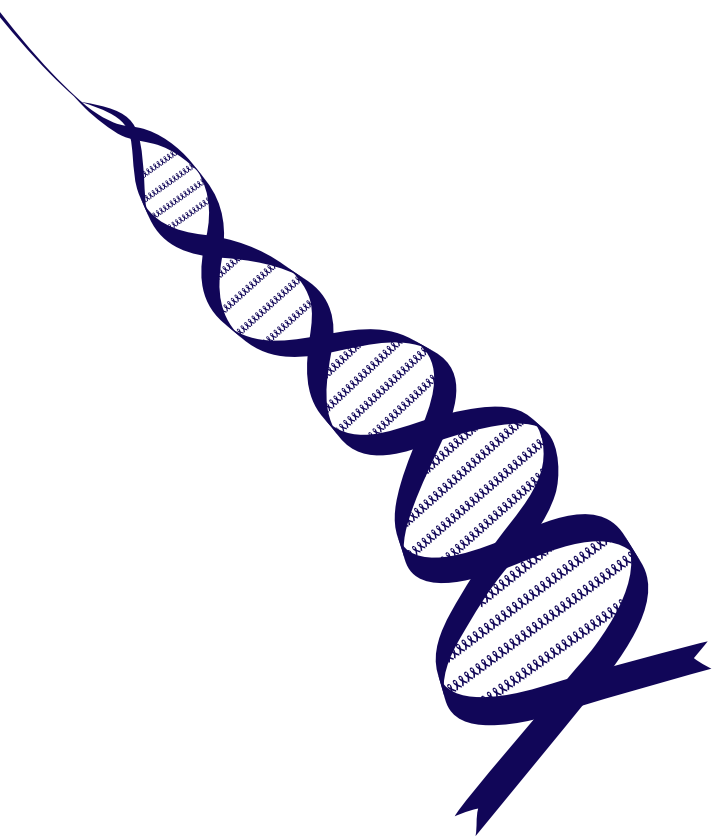


Chapter 6

ER pathway activity as a predictive marker during neo-adjuvant endocrine therapy in early breast cancer; Results of the TEAM IIA trial

E.J. Blok
M. Alves de Inda
P.J.K. Kuppen
A.C. Charehbili
E. den Biezen-Timmermans
S. Fruytier
W.M. Meershoek-Klein Kranenbarg
A. van Brussel
S.C. Linn
W. Verhaegh
A. van de Stolpe
C.J.H. van de Velde
J.R. Kroep

Manuscript as submitted to
Cancer Research



Abstract

Endocrine therapy is an important asset in the management of estrogen receptor (ER) positive breast cancer. Currently, patients are selected for endocrine therapy based on immunohistochemical expression of ER and progesterone receptor (PR). However, this does not necessarily imply an active ER pathway, which is the target of endocrine therapy. The aim of this study was to validate a recently described computational model that infers ER pathway activity based on the expression of its target genes.

Two cohorts were analyzed: a cohort of 61 patients treated with 3 months of neo-adjuvant letrozole for which a public dataset containing gene expression and ultrasound-based tumor-response data is available, and the TEAM-IIA trial cohort in which 102 patients were treated with up to 6 months of neo-adjuvant exemestane. The original ER pathway model was adapted to process Affymetrix HGU133A data from the public dataset, and qPCR data obtained from TEAM-IIA samples.

Mean ER pathway activity decreased significantly during therapy. Furthermore, in the public dataset, both baseline activity and decrease in ER pathway activity was significantly higher in responding patients. In the TEAM-IIA cohort, palpation-based progressive disease and radiology-based non-response after therapy were associated with lower levels of baseline ER pathway activity.

Our results indicate, in two independent cohorts, that low baseline ER pathway activity is associated to an inferior response to endocrine therapy. Upon prospective validation, this model could be used in a clinical setting to predict response to endocrine therapy and thereby better select patients who will benefit from this treatment.

Introduction

Endocrine therapy is one of the mainstays in the treatment of both early and metastatic breast cancer. Especially the use of tamoxifen and aromatase inhibitors (AI) has resulted in increased survival rates.[1-3] Patients are currently selected for endocrine therapy using immunohistochemical analysis of estrogen receptor (ER) and progesterone receptor (PR) expression, which was developed more than a decade ago.[4] Both the American Society for Clinical Oncology (ASCO) and the European Society for Medical Oncology advise a threshold of 1% ER positive tumor cells.[5, 6] In practice, many clinicians and countries choose a threshold of 10%.[7-9] More quantifiable analyses like the Allred scoring and H-score have been developed and suggested for clinical application, but are currently not routinely used.[5]

Despite the success of endocrine therapy in ER-positive breast cancer, there are still patients that do not respond to endocrine therapy, regardless of the presence of ER or PR in the investigated tissue sample. In addition to cancer tissue heterogeneity, several mechanisms have been proposed to explain lack of therapy response, like emergence of activating mutations in ESR1 or activation of other signal transduction pathways upon pharmacological inhibition of the ER pathway.[10, 11] Standard immunohistochemical analysis only detects presence of ER and PR; however, the issue of how well a positive nuclear ER staining, accompanied or not by a positive nuclear PR staining, actually indicates an active ER pathway, as an alternative explanation, has not yet been satisfactorily addressed. The development of tests to predict response to endocrine therapy based on measuring actual activity of the ER pathway will provide additional information that might be useful in the decision to treat with endocrine therapy and whether to add additional targeted therapy.

A possible approach towards predicting the response of endocrine therapy would be to assess activity of the ER pathway by measuring expression of downstream ER pathway target genes. After all, it is likely that if this pathway is highly active, endocrine therapy will be more effective than when the pathway is barely active or inactive, irrespective of the presence of the estrogen receptor itself. So far, no test has been developed for assessing ER pathway activity, although numerous studies have been performed.[12-15] Recently, Verhaegh *et al* have developed a computational model for the ER pathway, enabling assessment of this pathway in tumor tissue. This computational Bayesian network model uses mRNA expression of 27 genes which are proven target

genes of the ER pathway.[16, 17] The aim of the current work is to evaluate this model in a clinical setting. We hypothesize that ER pathway activity at baseline as measured by this computational model is capable of predicting response to therapy and that a decrease in pathway activity during treatment represents a successful treatment.

To test our hypothesis, we investigated publicly available Affymetrix datasets containing baseline and outcome data from neoadjuvant studies on AIs in breast cancer patients.[18, 19] After initial confirmation of our hypothesis we proceeded to test it in the TEAM IIA clinical trial cohort of ER positive breast cancer patients treated with neo-adjuvant endocrine therapy. We assessed the value of baseline ER pathway activity to predict response to neo-adjuvant therapy with an AI and whether the decrease in activity could be used to assess clinical response to neoadjuvant therapy.

Materials & Methods

Publicly available dataset

Search of the GEO repository yielded one dataset appropriate for our proof-of-principle analysis: the GSE20181 dataset.[18, 19] This dataset contains Affymetrix HGU133A gene expression and outcome data of 61 post-menopausal patients that underwent neo-adjuvant letrozole treatment. Those patients were treated in a neo-adjuvant setting with 2.5 mg daily letrozole for three months. The GSE20181 dataset contains gene expression data from biopsies, containing at least 20% tumor collected at baseline, 14 days, and three months of treatment. Clinical response was determined by ultrasound assessment of the tumor size and patients with more than 50% tumor reduction were considered responders. Raw Affymetrix data from the GSE20181 dataset was retrieved from the GEO repository and normalized with fRMA.[20]

Study cohort

The Dutch TEAM IIA trial is a neo-adjuvant phase II trial for which the details have been previously described.[21] Briefly, 102 ER positive patients were randomized for either 3 or 6 months of neo-adjuvant exemestane. Due to unforeseen slow accrual, the study changed to a single arm design with 6 months of therapy. Standard clinicopathological baseline characteristics, including PR and HER2 status were known. Pre-treatment biopsies and post-treatment resection specimens were collected by the investigators. Change in tumor size assessed by palpation was the primary objective. Secondary

outcomes were clinical response rates measured by mammography, ultrasound and MRI. Assessment was performed according to RECIST 1.1 criteria.

Laser Capture Microdissection

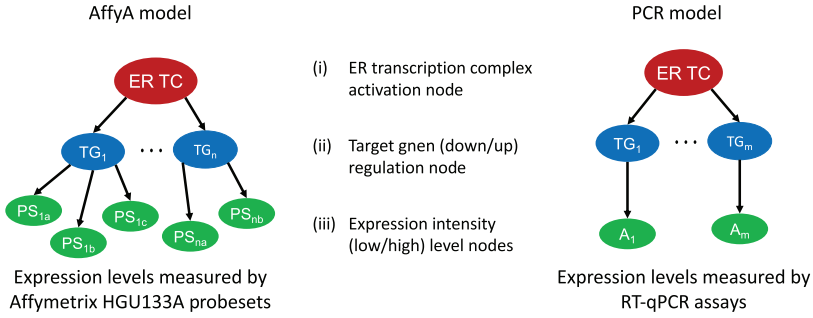
In order to accurately determine the activity of the ER pathway within tumor cells, those cells were separated from surrounding stromal tissue using Laser Capture Microdissection (LCM), as described by Espina *et al.*[22] Briefly, slides were cut in quintuplicate from tumor-containing formalin fixed paraffin embedded (FFPE) blocks from both pre-treatment biopsies as well as from post-treatment resection material. The slides were cut and placed onto PEN Membrane slides (Life Technologies), which contain a special membrane making it suitable for LCM. Slides were stained with hematoxylin and eosin using standard RNase free protocols. All slides were microscopically assessed and tumor regions (at least 1mm^2 per slide) were dissected from the slides using LCM. For each sample, the dissected tumor regions were collected in an RNase-free microfuge tube (Life Technologies) and were stored at -20 degrees Celsius.

ER pathway models

The methodology used to develop the ER pathway models used in this study was described in detail earlier.[16] The basic idea behind this methodology is to construct a Bayesian network model of the ER transcriptional program, which uses the pathway target genes' mRNA levels in a certain sample to infer the probability that the ER pathway is transcriptionally active in the respective sample. This Bayesian network structure is a simplified model of the transcriptional program of the ER signal transduction pathway, consisting of three types of nodes: (i) ER transcription complex activation node, (ii) target gene regulation node (with states 'down' and 'up'), and (iii) expression intensity level nodes (with states 'low' and 'high') each corresponding to a target gene. The final model describes (a) how target gene regulation depends on ER transcription complex activity and (b) how expression level intensities in turn depend on regulation of the respective target genes (Supplemental Figure 1).

The original paper described this methodology for mRNA expression levels measured using the Affymetrix HGU133Plus2.0 microarray platform.16 This methodology was adapted to be used with both RT-qPCR and Affymetrix HGU133A platforms. This adaptation consisted of modifying and re-calibrating the ER pathway Bayesian network model for each platform. For the RT-qPCR platform the original 27 target gene ER

pathway Bayesian network was reduced to a 12 ER target gene network. The genes included in this network were chosen based on literature evidence and discriminative power. RT-qPCR assays were developed and validated for each of these target genes plus 7 other genes used as reference genes. The model was calibrated using mRNA expression data of MCF7 cell cultures exposed to 1nM E2 for 16h or DMSO after being maintained in estradiol deprived (charcoal treated FBS, phenol red free) medium for 48h, as described before.[23] For the Affymetrix HGU133A platform the network was adapted by selecting Affymetrix probe sets representing the original 27 ER target genes that were available in the HGU133A platform and the resulting network model was calibrated using fRMA transformed Affymetrix HGU133A mRNA expression data from publicly available dataset GSE9936. These Affymetrix HGU133A calibration samples were from MCF7 cell cultures exposed to 6nM E2 for 24h or vehicle control after being maintained in estradiol deprived (charcoal treated FBS, phenol red free) medium for at least 48h.[24]



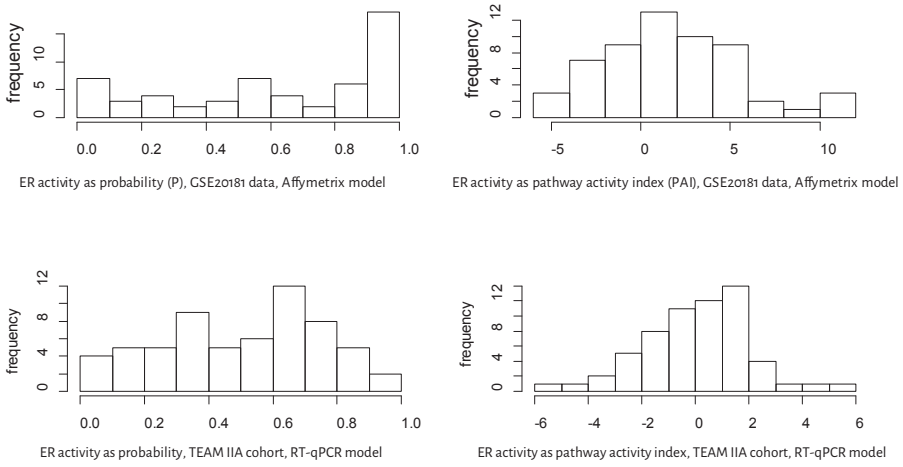
Suppl Figure 1: schematic representation of ER pathway network. Left: RT-qPCR model consisting of 12 target genes with expression levels measured by RT-qPCR assays. Center: Node type description. Right: Affimetrix model consisting of 27 target genes with expression levels measured by Affymetrix hgu133a probesets.

Pathway analysis of PCR samples

For the analysis of the ER pathway activity on the samples of the TEAM IIA cohort, mRNA was extracted from the micro dissected samples (Siemens VERSANT Tissue Prep. Reagent kit, according to manufacturer's instructions) and the expression of the 12 target genes and 7 reference genes was determined using RT-qPCR. Those values were used as input to the RT-qPCR ER pathway model with which the probability and the odds that the ER pathway is active was calculated. These data were interpreted with respect to ER pathway activity in a blinded manner at Philips Research and returned to the Leiden University Medical Center, in order to be correlated to tumor size, as assessed by palpation and mammography.

Statistical analysis

Statistical analysis was performed using R[25] and SPSS 23.0 (IBM). ER pathway analysis was calculated in two ways: the probability (P) that the ER pathway is active and its *pathway activity index* (PAI, defined as $\log_2(P/(1-P))$, i.e., the $\log_2$ of the odds in favor that the ER pathway is active, PAI=0 indicates a 50% chance of ER pathway activation, PAI=2 and -1 indicate, respectively, a chance four times as large and twice as small of the ER pathway being active than of it being inactive). For the purpose of comprehensibility, the probability is used as a visual representation of the likelihood that the ER pathway is active. However, the activity index is used for all statistical calculations since these data are generally more appropriate for statistical computations and more normal-like distributed (Supplemental Figure 2). For simplicity, we use 'ER pathway activity', or simply, 'ER activity' as a short form of 'the inferred probability that the ER pathway is active'. Paired t-tests were used to assess differences between baseline and post-treatment ER pathway activity. Two sample t-tests were used to assess correlations between outcome categories and average ER pathway activity. Two sample t-tests and ANOVA were used to access ER activity and decrease in activity association with baseline parameters. Presented p-values are 2 sided and refer to PAI quantities, unless explicitly indicated otherwise. All tests assume unequal variance.



Suppl Figure 2: histograms of ER-pathway activity at baseline. Upper: GSE20181 data, using Affymetrix model; Lower: TEAM II A cohort, using RT-qPCR model; Left: activity presented as a probability (P); Right: activity presented as Pathway Activity Index, i.e.: $\log_2(P/(1-P))$.

Results

GSE20181 proof-of-principle dataset

We first assessed, in a public dataset, whether the ER pathway model based on Affymetrix HGU133A microarray data, was able to detect a decrease in ER pathway activity in ER positive patients following neoadjuvant endocrine therapy. ER pathway activity significantly decreased during letrozole therapy on ER positive patients of the GSE20181 dataset (table 1) presents the mean ER pathway activity at baseline, 2 weeks, and 3 months. Overall, ER pathway mean activity decreased significantly between baseline (P=0.62) and 2 weeks (P=0.32) of treatment (with PAI decrease from 1.6 to -1.6, paired t-test p-value <0.001, n=58) as well as between baseline and 3 months (P=0.36) of therapy (with PAI decrease from 1.8 to -1.3, p-value <0.001, n=56). No significant difference was seen between 2 weeks and 3 months of treatment, suggesting an early maintained response to letrozole. Both baseline ER pathway activity and decrease in ER activity from baseline was significantly higher in responders than in non-responders (Table 1 and Figure 1).

Table 1: Average activity and decrease in activity of ER pathway for GSE20181. ER pathway activity is presented as probability that the ER pathway is active (P) and as activity index (PAI). R: responders, NR: non-responders;; sd: standard deviation; pv: p-value for t-test on PAI for R vs. NR

	P				PAI			
	N	Mean	sd	Range	Mean	sd	Range	p-value*
Baseline								0.02
all	58	0.62	0.33	[0.02,1]	1.6	3.8	[-5.8,12]	
R	37	0.65	0.34	[0.02,1]	1.9	3.9	[-5.8,10]	
NR	15	0.46	0.31	[0.06,0.92]	-0.3	2.4	[-3.9,3.5]	
2 weeks								0.86
all	58	0.32	0.27	[0.01,0.96]	-1.6	2.6	[-6.7,4.5]	
R	37	0.33	0.29	[0.01,0.96]	-1.6	2.9	[-6.7,4.5]	
NR	15	0.32	0.23	[0.01,0.76]	-1.5	2	[-6.2,1.7]	
3 months								0.86
all	60	0.38	0.28	[0.01,1]	-1.1	2.6	[-6.6,8.2]	
R	36	0.36	0.26	[0.03,0.86]	-1.2	2	[-4.8,2.7]	
NR	14	0.40	0.33	[0.01,0.91]	-1.0	3	[-6.4,3.4]	
Base - 2 w								0.02
all	58	0.30	0.35	[-0.56,0.93]	3.2	3.5	[-3.9,10]	
R	37	0.31	0.33	[-0.33,0.93]	3.5	3.5	[-3.1,10]	
NR	15	0.14	0.35	[-0.56,0.68]	1.2	2.7	[-3.9,5.1]	

Table 1: continued

		P			PAI			
	N	Mean	sd	Range	Mean	sd	Range	p-value*
Base - 3 m								0.04
all	56	0.28	0.37	[-0.8,0.89]	3.1	3.8	[-6.4,12]	
R	36	0.31	0.31	[-0.46,0.88]	3.3	3.4	[-4.3,10]	
NR	14	0.07	0.43	[-0.8,0.82]	0.76	3.8	[-6.4,8.6]	
2 w - 3 m								0.96
all	56	-0.03	0.23	[-0.66,0.41]	-0.29	2.2	[-5.1,3.8]	
R	36	-0.02	0.23	[-0.63,0.41]	-0.35	2.1	[-5.1,3.7]	
NR	14	-0.07	0.25	[-0.66,0.23]	-0.39	2.5	[-5,3.8]	

*p-value represents a t-test comparing responders (R) with non-responders (NR)

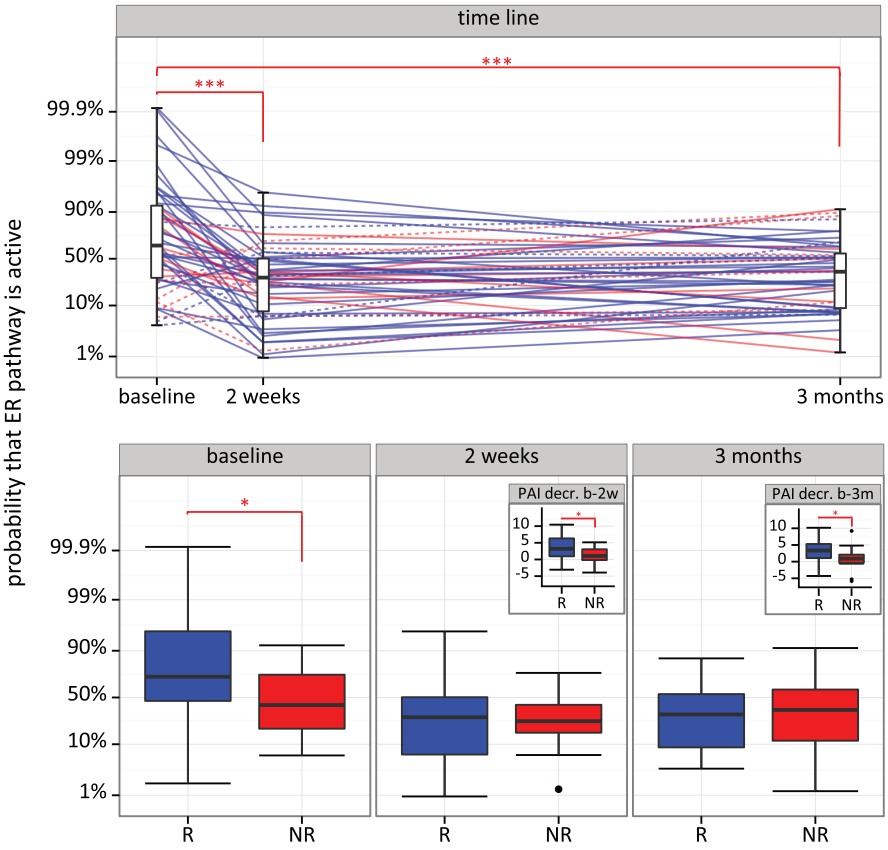


Figure 1: ER-pathway activity as a function of treatment time for GSE20181 dataset. Upper: line plot of activity per sample. Lower: box plots of activity in responders and non-responders per time point. Blue: responders, red: non-responders. Solid lines: activity went down during treatment, dashed lines: activity went up during treatment.

Subsequently we analyzed the tissue samples from the TEAM IIA clinical trial, using the RT-qPCR ER pathway model.

TEAM IIA cohort characteristics

The TEAM IIA cohort consists of 102 patients, all of which received at least 3 months of neoadjuvant endocrine therapy with exemestane. The majority of patients (n=83) received 6 months of therapy. Both biopsies and resection samples were collected and analyzed retrospectively. Thirty-three biopsies and 38 resection samples could not be retrieved, leading to a cohort of 69 patients from whom a biopsy sample was available and a cohort of 64 patients from whom a resection specimen was available. During LCM another 16 resection samples and 11 biopsies were excluded, mainly due to absence of tumor tissue in the specimen. Upon subsequent RT-qPCR, 9 biopsy samples yielded amounts of RNA too low to perform the required set of qPCRs, making them ineligible for further analysis, resulting in 49 biopsy and 48 resection samples eligible for analysis, of which 28 samples were matched cases from the same patient (Figure 2).

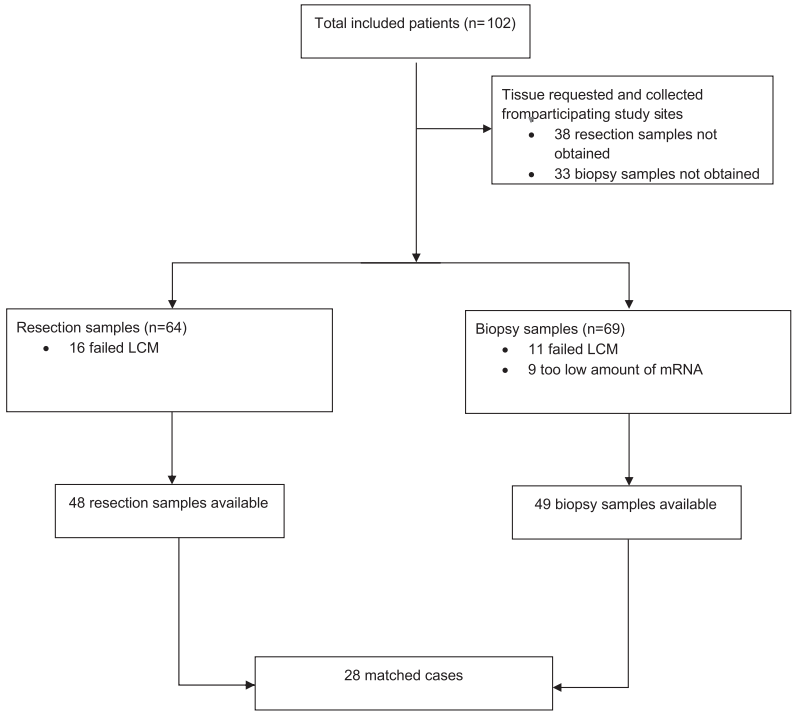


Figure 2 : consort diagram of tissue amples collected in the TEAM IIA trial

Baseline and post treatment ER pathway activity TEAM IIA

At baseline, the average probability of the ER pathway being active was 0.48, ranging from 0.02 to 0.89. After therapy, the average ER activity decreased to 0.32, ranging from 0.01 to 0.86 (Table 2). Of the 28 matched cases, the mean activity of ER decreased by 35% (absolute difference of -0.16), ranging from a decrease of 0.70 to an increase of 0.31 (mean PAI decrease from -0.3 to -1.8, paired t-test p-value <0.001).

ER pathway activity at baseline and decrease in activity during therapy were correlated with clinicopathological baseline characteristics (Table 3). Baseline mean ER pathway activity was significantly higher in PR positive samples (P=0.55) than PR negative samples (P=0.29), with PAI=0.3 vs. -1.8, two sample t-test p-value=0.003; and in patients with higher BMI (P=0.53) than in patients with lower BMI (P=0.38), with PAI=0.2 vs. -1, p-value=0.04. There was no significant difference in baseline activity between age categories, HER2-status, tumor type, tumor grade and tumor or nodal status.

Interestingly, the mean decrease in ER pathway activity in the available paired cases was neither significantly higher in baseline PR postive patients than in baseline PR negativepatients nor in patients with higher baseline BMI than in patients with lower baseline BMI (Table 3).

Correlation with TEAM IIA outcome assessed by palpation

To test the predictive value of the ER pathway model, the computed ER pathway activities were correlated to primary and secondary outcome measures of the TEAM IIA trial. The primary outcome of the trial was response to therapy based on palpation. Analysis showed that, at the end of therapy, patients with a palpation-based progressive disease (PD) had a significant lower ER pathway activity at baseline (P=0.16) compared to patients with complete remission (CR; P=0.59), partial remission (PR; P=0.48) or stable disease (SD; P=0.53); mean PAI PD=-2.5 vs non-PD=0.1, p-value=0.03 (Figure 3). Similar prognostic effects of ER pathway activity were shown at three months, as a measure of early response; p-value=0.038.

Table 2: ER pathway at baseline and resection for the TEAMII A cohort. ER pathway activity is presented as probability that the ER pathway is active (P) and as pathway activity index (PAI). CR: complete remission, PR: partial remission, SD: stable disease, PD: progressive disease; av: average, sd: standard deviation.

	Baseline							Resection						
	N	P			PAI			N	P			PAI		
		Mean	sd	Range	Mean	sd	Range		Mean	sd	Range	Mean	sd	Range
Total	49	0.48	0.25	[0.02, 0.89]	-0.26	1.9	[-5.6, 3.1]	50	0.31	0.21	[0.01, 0.86]	-1.65	2.3	[-6.3, 2.7]
Stratified by response assessed by palpation at last measurement														
CR	11	0.59	0.19	[0.36, 0.89]	0.63	1.3	[-0.8, 3]	9	0.20	0.18	[0.03, 0.57]	-2.5	1.7	[-5,0.38]
PR	12	0.48	0.28	[0.06, 0.81]	-0.32	2.1	[-4.1,2.1]	15	0.30	0.30	[0.01,0.86]	-2.0	2.7	[-6.3, 2.6]
SD	12	0.52	0.23	[0.11,0.80]	0.13	1.5	[-3,2]	13	0.40	0.23	[0.06, 0.77]	-0.8	1.7	[-3.9,1.8]
PD	3	0.16	0.11	[0.09, 0.29]	-2.50	1.1	[-3.3.-1.3]	3	0.08	0.05	[0.03, 0.12]	-3.8	1.2	[-5.2,-2.9]
Stratified by response assessed by mammography at last measurement														
CR/PR	0/11	0.50	0.31	[0.07,0.89]	-0.12	2.3	[-3.7,3]	1/9	0.29	0.26	[0.01,0.86]	-1.8	2.4	[-6.3, 2.6]
SD/PD	12/1	0.47	0.25	[0.6,0.8]	-0.28	1.8	[-4.1,2]	11/1	0.28	0.26	[0.3,0.73]	-2.0	2.2	[-5.2,1.4]

Table 3 Association of baseline parameters of the TEAM II-A cohort with ER pathway activity at biopsy and its decrease after therapy (difference between baseline biopsy and post-therapy resection sample; positive values represents a decrease in activity during therapy). ER pathway activity is presented as probability that ER is active (P) and as pathway activity index (PAI). %*: percentage excluding missing; sd: standard deviation; p-value ANOVA test

Total N=102			Baseline N=49							Decrease baseline - resection N=28						
	N	(%*)	N	(%*)	P Mean	Range	PAI Mean	Range	p-value	N	(%*)	P Mean	Range	PAI Mean	[range]	p-value
Histological grade									0.43							0.009
G1	12	(11.8)	8	(16.3)	0.37	[0.07,0.71]	-1.0	[-3.7,1.3]		2	(7.1)	0.21	[0.18,0.25]	2.9	[1.5,4.2]	
G2	27	(26.5)	13	(26.5)	0.50	[0.18,0.76]	0.0	[-2.2,1.6]		11	(39.3)	0.20	[-0.31,0.51]	0.0	[-2.2,1.5]	
G3	7	(6.9)	2	(4.1)	0.24	[0.11,0.37]	-1.9	[-3,-0.8]		2	(7.1)	-0.22	[-0.25,-0.20]	-1.9	[-3,-0.8]	
Gx	56	(54.9)	26	(53.1)	0.52	[0.02,0.89]	0.0	[-5.6,3]		13	(46.4)	0.20	[-0.20,0.70]	-0.3	[-4.1,2.4]	

Clinical nodal status								0.81						0.92
No	77	(75.5)	34	(69.4)	0.48	[0.02,0.84]	-0.3	[-5.6,2.4]	21	(75)	0.17	[-0.2,0.7]	1.5	[-1.9,5.2]
N1-3	25	(24.5)	15	(30.6)	0.48	[0.07,0.89]	-0.2	[-3.7,3]	7	(25)	0.14	[-0.31,0.47]	1.6	[-1.2,4.3]
Age category								0.32						0.45
50-59 y	8	(7.8)	4	(8.2)	0.27	[0.06,0.66]	-1.9	[-4.1,1]	3	(10.7)	0.14	[0,0.4]	1.2	[0,2.4]
60-69 y	32	(31.4)	13	(26.5)	0.47	[0.02,0.8]	-0.4	[-5.6,2]	5	(17.9)	0.01	[-0.25,0.51]	0.3	[-1.5,3.9]
≥ 70 y	62	(60.8)	32	(65.3)	0.51	[0.07,0.89]	0.0	[-3.7,3]	20	(71.4)	0.21	[-0.31,0.7]	1.9	[-1.9,5.2]
Tumor type								0.19						**
Ductal	66	(66.7)	34	(72.3)	0.50	[0.06,0.89]	-0.1	[-4.1,3]	21	(77.8)	0.18	[-0.31,0.7]	1.5	[-1.9,5.2]
Lobular	30	(30.3)	11	(23.4)	0.41	[0.02,0.8]	-0.9	[-5.6,2]	5	(18.5)	0.17	[0.03,0.34]	2.3	[0.2,4.3]
Other	3	(3.0)	2	(4.3)	0.72	[0.63,0.81]	1.4	[0.8,2.1]	1	(3.7)	0.16	-	1.2	-
PR status								0.003						0.42
Negative	32	(31.4)	13	(26.5)	0.29	[0.02,0.63]	-1.8	[-5.6,0.8]	6	(21.4)	0.06	[-0.2,0.47]	0.8	[-1.8,4.3]
Positive	70	(68.6)	36	(73.5)	0.55	[0.09,0.89]	0.3	[-3.3,3]	22	(78.6)	0.19	[-0.31,0.7]	1.7	[-1.9,5.2]
HER2 (IHC/FISH)								0.67						**
Negative	85	(88.5)	45	(91.8)	0.49	[0.02,0.89]	-0.2	[-5.6,3]	26	(92.9)	0.18	[-0.31,0.66]	1.7	[-1.9,5.2]
Positive	9	(9.4)	2	(4.1)	0.26	[0.11,0.41]	-1.8	[-3,-0.5]	1	(3.6)	-0.2	-	-1.8	-
Undetermined	2	(2.1)	2	(4.1)	0.46	[0.25,0.68]	-0.3	[-1.6,1.1]	1	(3.6)	0.03	-	0.2	-
Tumor size								0.97						0.41
< 3 cm	43	(43.4)	21	(44.7)	0.48	[0.02,0.89]	-0.3	[-5.6,3]	9	(33.3)	0.08	[-0.31,0.66]	0.7	[-1.9,4.6]
3-5 cm	41	(41.4)	21	(44.7)	0.48	[0.07,0.84]	-0.2	[-3.7,2.4]	16	(59.3)	0.23	[-0.2,0.7]	2.0	[-1.5,5.2]
> 5 cm	15	(15.2)	5	(10.6)	0.48	[0.19,0.8]	-0.1	[-2.1,2]	2	(7.4)	0.22	[0.18,0.27]	2.9	[1.6,4.2]
BMI								0.04						0.46
< 25	33	(32.4)	19	(38.8)	0.39	[0.02,0.81]	-1.0	[-5.6,2.1]	12	(42.9)	0.09	[-0.31,0.7]	1.2	[-1.9,5.2]
> 25	69	(67.6)	30	(61.2)	0.54	[0.09,0.89]	0.2	[-3.3,3]	16	(57.1)	0.22	[-0.25,0.66]	1.8	[-1.5,4.6]

** not enough sample to perform statistical analysis; ≠mean differs from other categories

ER pathway activity decreased in all six patients with complete remission, while 29% (two of seven) patients with partial remission and 38% (three of eight) of patients with stable disease showed an increase in ER activity. Most notably, the decrease in ER pathway activity was significantly higher in the combined CR/PR/SD group than in the PD group, whereas ER activity did not change significantly in the two patients with progressive disease (mean decrease in probability = 0.18 vs -0.002; PAI = -0.1 vs -3.2, paired t-test p-value=0.003) (Table 4). These observations are in accordance with the hypothesis that hormone therapy is only effective in patients with an active ER pathway and that success of response is associated with a decrease in ER pathway activity during treatment.

Table 4: Welsh two samples t-test for mean ER activity and mean decrease in ER activity during treatment for the TEAMIA cohort. ER pathway activity is presented as probability that the ER pathway is active (P) and as pathway activity index (PAI). Evaluation is performed on baseline biopsies, post-treatment resection samples, and the difference between those timepoints. The response is evaluated by palpation and mammography at 3 months and at last therapy. CR: complete remission, PR: partial remission, SD: stable disease, PD: progressive disease; p-values were computed based on PAI data.

	Group	N	P mean (sd)	PAI mean (sd)	Group	N	P mean (sd)	PAI mean (sd)	p-value
Stratified by response assessed by palpation at 3 months									
Baseline	CR/PR/SD	27	0.53 (0.23)	0.14 (1.7)	PD	3	0.17 (0.12)	-2.5 (1.2)	0.04
Stratified by response assessed by palpation at last measurement									
Baseline	CR/PR/SD	35	0.53 (0.11)	0.13 (1.7)	PD	3	0.16 (0.11)	-2.5 (1.1)	0.03
Resection	CR/PR/SD	37	0.31 (0.26)	-1.7 (2.2)	PD	3	0.08 (0.5)	-3.8 (1.2)	0.07
Baseline-Resection	CR/PR/SD	21	0.18 (2.28)	1.6 (2.2)	PD	2	-0.002 (0.003)	-0.02 (0.05)	0.003
Stratified by response assessed by mammography at 3 months									
Baseline	CR/PR	6	0.71 (0.17)	1.5 (1.2)	SD/PD	12	0.44 (0.24)	-0.49 (1.7)	0.015
Stratified by response assessed by mammography at last measurement									
Baseline	CR/PR	11	0.5 (0.31)	-0.12 (2.3)	SD/PD	13	0.48 (0.25)	-0.28 (1.8)	0.86
Resection	CR/PR	10	0.29 (0.26)	-1.8 (2.4)	SD/PD	12	0.28 (0.26)	-2 (2.2)	0.85
Baseline-Resection	CR/PR	6	0.18 (0.21	1.7 (2.4)	SD/PD	7	0.21 (0.34)	1.7 (2.4)	0.99

Correlation with TEAM IIA outcome assessed by mammography

Since response based on mammography served as a secondary outcome in the TEAM IIA trial, mammography was not mandatory for every patient. Due to the lower number of available observations in the CR and PD categories, CR and PR were combined into a category of responders, whereas SD and PD were combined into a category of non-responders. At three months, responders and non-responders could be clearly separated, based on the ER activity at baseline (responders P=0.71 vs non-responders P=0.44; PAI=1.5 vs 0.5, two samples t-test p-value=0.015). At 6 months

however, no distinction could be made (responders P=0.5 vs non-responders P=0.48; PAI=-0.1 vs -0.2, p-value=0.9).

Clinical validation

In summary, we performed an ER pathway analysis in two distinct AI neoadjuvant settings: the first, a proof of principle, public Affymetrix dataset derived from a cohort of patients undergoing letrozole AI therapy for three months; the second, a trial cohort of patients that underwent 3 to 6 months exemestane AI therapy. Both cohorts where assessed using adapted ER pathway models based on the original Affymetrix ER pathway model [16]. While the proof-of-concept dataset analysis used Affymetrix HGU133A gene expression data from fresh-frozen tissue, the TEAM IIA trial cohort analysis used RT-qPCR data from FFPE tissue obtained by LCM. Furthermore the two models used a different number target genes (27 vs. 12) and were calibrated using data from MCF7 cell cultures stimulated different estradiol concentration (6 nM vs. 1 nM).

Despite the differences between cohorts and methods, both analyses detected a significant decrease in ER pathway activity during AI treatment and indicated that the decrease in activity, as well as baseline ER pathway activity were significantly higher in responders than in non-responders, though response was assessed differently in the two studies. The proof-of-concept study defined response as a reduction in tumor size of at least 50% measured by ultrasound, while the TEAMII A cohort adopted the RECIST 1.1 criteria (at least 30% tumor reduction) using response measured by palpation as primary outcome and measured using radiological modalities as secondary outcomes.

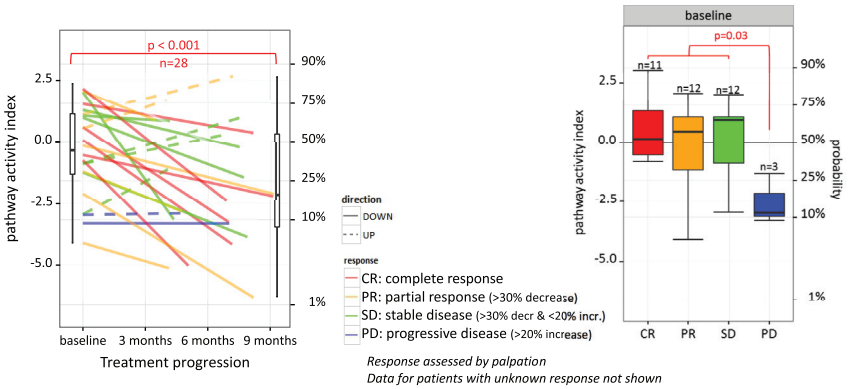


Figure 3: ER-pathway activity in patients of the TEAMIIA cohort stratified by response assessed by palpation at last measurement. Left: Line plots of ER-pathway activity at baseline and resection. Right: box plots of ER activity at baseline. Red: complete remission, orange: partial remission, green: stable disease, blue: progressive disease. Solid lines: activity went down during treatment, dashed lines activity went up during treatment.

Discussion

Duration of treatment and resistance

In the GSE20181 cohort, a significant difference in ER pathway activity between baseline and both two weeks and three months of letrozole therapy, but not between two weeks and three months, indicating an early and maintained response. A remarkable finding is the observation that baseline ER pathway activity was predictive for therapy response as measured by mammography at three months, but not at six months of therapy. This difference may be explained by mechanisms of treatment resistance. During the course of endocrine therapy, over 20% of breast cancers are known to acquire resistance against aromatase inhibitor treatment due to activating mutations in the estrogen receptor, resulting in reactivation of the ER pathway driving tumor growth. Alternatively other signal transduction pathways, like the PI3K pathway may take the lead in driving tumor growth.[11]

Another reason for loss of treatment effect during prolonged therapy, could be lower compliance at the end of therapy, resulting in reactivation of the ER pathway and the subsequent regrowth of the tumor. Although data on compliance are not available in this trial, it seems unlikely that non-compliance has occurred since patients were monitored closely during the trial.

Correlation with baseline parameters

The significantly higher baseline ER pathway activity in PR positive cases in the TEAM IIA cohort was expected, since PGR, the gene that codes for PR, is a target gene of the ER pathway. However, the decrease in ER pathway activity was not significantly higher in PR positive cases and the lack of correlation between baseline PR status and therapy response indicates that PR status by itself it is not sufficient to accurately infer pathway activation, or its deactivation in connection with AI therapy.

Study limitations

Drawbacks of this study are the retrospective setting, the limited availability of sample with sufficient quality, and the lack of a reliable read-out for the primary and secondary clinical outcome, that is, reduction in tumor size. Although the entire TEAM IIA cohort comprises 102 patients, due to described logistical reasons only 28 paired cases were available for analysis, limiting statistical power. Palpation was the primary outcome in this trial, whereas tumor reduction assessed by radiological imaging modalities was a secondary outcome, and therefore not performed in all patients. In addition, it is generally accepted that radiological evaluation using MRI in particular, may not be optimal in terms of accuracy for evaluation of therapy response in neo-adjuvant treatment of ER positive breast cancer.[26, 27] In this specific cohort, a comparison of tumor size at resection versus its estimated size by palpation, mammography, US, and MRI suggested that palpation and mammography were the most accurate methods, while in many cases, various methods provided discordant results.[28] Therefore only evaluations based on palpation and mammography were included in this analysis.

Another point of concern is tumor heterogeneity. ER activity as measured in the biopsy setting may not be representative of the whole tumor, which could be a source of error in the analysis, and could explain why some patients with low baseline ER activity showed some response to treatment.

Future perspectives

Having comparable results from two distinct ER positive breast cancer cohorts, a validated Affymetrix gene expression dataset and a trial cohort with daily practice FFPE tissue samples for mRNA qPCR analysis, this study offers a robust clinical validation of the ER pathway model, developed by Verhaegh *et al*[16], both when the full spectrum of target genes is measured using Affymetrix microarray on fresh tissue,

and when the selected subset of target genes is measured using multiple qPCR assays on FFPE material.

Further prospective validation of the ER pathway model is necessary before this approach can be used in a routine clinical setting to predict endocrine therapy response. A prospective study to validate the use of the ER pathway test to predict response to neo-adjuvant endocrine therapy is being planned as a side study in a CDK4/6 inhibitor clinical trial. This side study will investigate to which extent the ER pathway is capable of defining which patient benefits most from endocrine therapy alone, which patient should be offered a combination therapy with targeted agents such as CDK4/6 inhibitors, or which patient should be offered chemotherapy directly.

Upon extended clinical validation, our findings could lead to a new diagnostic test to quantify ER pathway activity in a cancer tissue sample, with clearly defined clinical utility, to be implemented in stratification of ER positive breast cancer patients. For example, some studies suggest that neo-adjuvant hormonal therapy is a good alternative for chemotherapy for specific forms of breast cancer.[29, 30] Measuring ER pathway activity could be used to help decide on treatment course. Low baseline ER pathway activity wouldindicatea low chance of response to hormonal therapy, in which case chemotherapy could be indicated. In addition, in case hormonal therapy is selected, a second biopsy after one month of therapy to determine the remaining ER pathway activity could be implemented to assess therapy response. Persistent high levels of ER pathway activity could indicate resistance to therapy or non-compliance. This would lead to a tailored neo-adjuvant treatment with improved response monitoring, and a stepping stone towards personalized management of breast cancer.

References

1. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
2. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. The Lancet 2009;381(9869):805-816.
3. Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.
4. Harvey JM, Clark GM, Osborne CK et al. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol 1999;17:1474-1481.
5. Hammond ME, Hayes DF, Dowsett M et al. American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28:2784-2795.
6. Senkus E, Kyriakides S, Ohno S et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2015;26(suppl 5):v8-v30.
7. NABON. Richtlijn Mammacarcinoom. 2012. 23-12-2015. Ref Type: Internet Communication
8. Yi M, Huo L, Koenig KB et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. Ann Oncol 2014;25(5):1004-1011.
9. Harbeck N, Rody A. Lost in Translation? Estrogen Receptor Status and Endocrine Responsiveness in Breast Cancer. JCO 2012;30(7):686-689.
10. Jeselsohn R, Buchwalter G, De Angelis C et al. ESR1 mutations[mdash]a mechanism for acquired endocrine resistance in breast cancer. Nat Rev Clin Oncol 2015;12(10):573-583.
11. Ciruelos Gil EM. Targeting the PI3K/AKT/mTOR pathway in estrogen receptor-positive breast cancer. Cancer Treatment Reviews 2014;40(7):862-871.
12. Renoir JM, Marsaud V+, Lazennec G. Estrogen receptor signaling as a target for novel breast cancer therapeutics. Biochemical Pharmacology 2013;85(4):449-465.
13. Gross JM, Yee D. How does the estrogen receptor work? Breast Cancer Res 2002;4(2):62-64.
14. Pietras RJ, Marquez-Garban DC. Membrane-Associated Estrogen Receptor Signaling Pathways in Human Cancers. Clinical Cancer Research 2007;13(16):4672-4676.
15. Frasor J, Danes JM, Komm B et al. Profiling of Estrogen Up- and Down-Regulated Gene Expression in Human Breast Cancer Cells: Insights into Gene Networks and Pathways Underlying Estrogenic Control of Proliferation and Cell Phenotype. Endocrinology 2003;144(10):4562-4574.
16. Verhaegh W, van Ooijen H, Inda MA et al. Selection of Personalized Patient Therapy through the Use of Knowledge-Based Computational Models That Identify Tumor-Driving Signal Transduction Pathways. Cancer Research 2014;74(11):2936-2945.
17. Verhaegh W, de Stolpe AV. Knowledge-based computational models. Oncotarget; Vol 5, No 14 2014.
18. Miller WR, Larionov A. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. Breast Cancer Research 2010;12(4):1-9.
19. Miller WR, Larionov A, Anderson TJ et al. Sequential changes in gene expression profiles in breast cancers during treatment with the aromatase inhibitor, letrozole. Pharmacogenomics J 2012;12(1):10-21.
20. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics 2010;11(2):242-253.
21. Fontein DBY, Charehbili A, Nortier JWR et al. Efficacy of six month neoadjuvant endocrine therapy in postmenopausal, hormone receptor-positive breast cancer patients: a phase II trial. European Journal of Cancer 2014;50(13):2190-2200.
22. Espina V, Wulfkuhle JD, Calvert VS et al. Laser-capture microdissection. Nat Protocols 2006;1(2):586-603.
23. Van de Wiel P, Verhaegh W, Alves de Inda M et al. 514: Assessing functional ER pathway activity using a computational pathway model. European Journal of Cancer50:S124.

24. Chang EC, Charn TH, Park SH et al. Estrogen Receptors a and b as Determinants of Gene Expression: Influence of Ligand, Dose, and Chromatin Binding. Molecular Endocrinology 2008;22(5):1032-1043.

25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. 2014.

Ref Type: Internet Communication

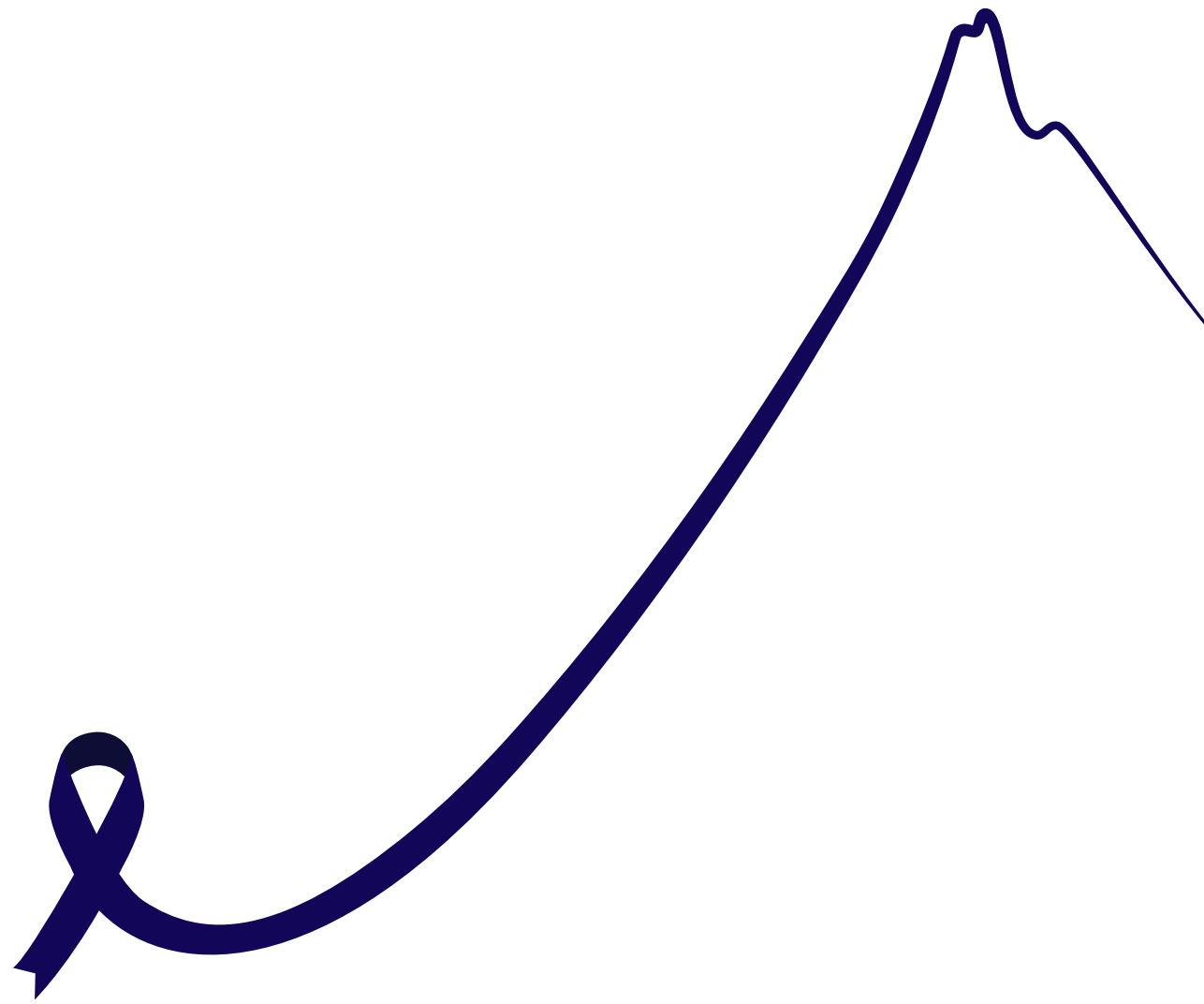
26. Charehbili A, Wasser MN, Smit VTHB et al. Accuracy of MRI for treatment response assessment after taxane- and anthracycline-based neoadjuvant chemotherapy in HER2-negative breast cancer. European Journal of Surgical Oncology (EJSO) 2014;40(10):1216-1221.

27. Chen JH, Feig B, Agrawal G et al. MRI evaluation of pathologically complete response and residual tumors in breast cancer after neoadjuvant chemotherapy. Cancer 2008;112(1):17-26.

28. Blok EJ, Charehbili A, Kroep JR et al. Abstract P6-01-05: Radiological evaluation of neo-adjuvant endocrine therapy in hormone-receptor positive early breast cancer. Cancer Research 2016;76(4 Supplement):6-01.

29. Purushotham A, Pinder S, Cariati M et al. Neoadjuvant Chemotherapy: Not the Best Option in Estrogen Receptor-Positive, HER2-Negative, Invasive Classical Lobular Carcinoma of the Breast? JCO 2010;28(22):3552-3554.

30. Dixon JM, Renshaw L, Dixon J et al. Invasive lobular carcinoma: response to neoadjuvant letrozole therapy. Breast Cancer Res Treat 2011;130(3):871-877.

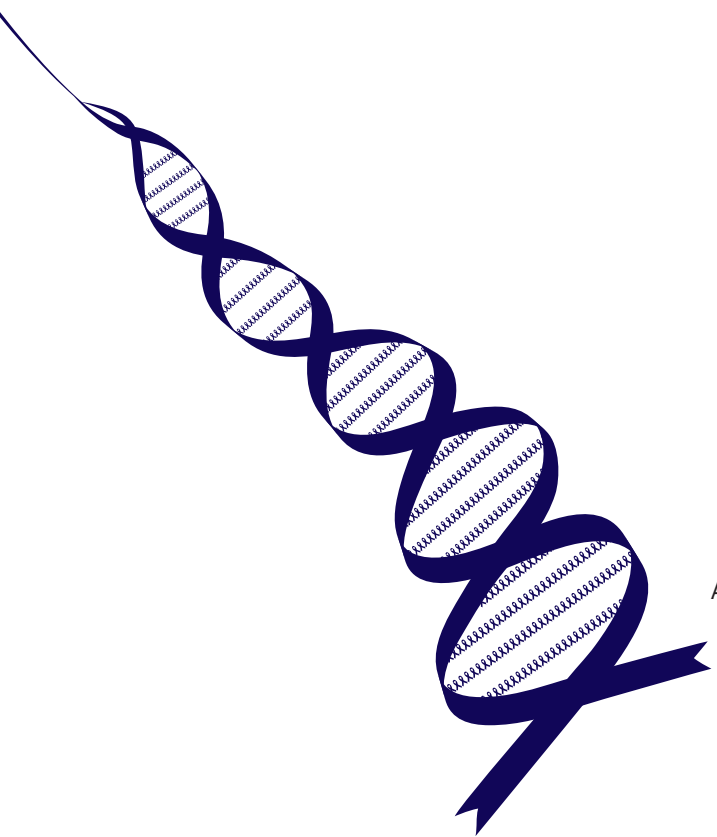


Chapter 7

Exploration of tumour-infiltrating lymphocytes as a predictive biomarker for adjuvant endocrine therapy in early breast cancer

E.J. Blok
C.C. Engels
N.G. Dekker-Ensink
W.M. Meershoek-Klein Kranenbarg
H. Putter
V.T.H.B.M. Smit
G.J. Liefers
J.P. Morden
J.M. Bliss
R.C. Coombes
J.M.S. Bartlett
J.R. Kroep
C.J.H. van de Velde
P.J.K. Kuppen

Accepted in: Breast Cancer Research and Treatment, 2018



Abstract

Background Tumour-infiltrating lymphocytes (TILs) have been shown to be prognostic for disease-free survival and predictive for the benefit of chemotherapy in patients with early breast cancer. The current study was performed to assess the predictive value of the number of CD8-positive TILs for the benefit of endocrine therapy with either tamoxifen or exemestane in two independent trial-cohorts.

Methods The number of CD8-positive TILs was assessed in a cohort of 236 Dutch breast cancer patients in the Intergroup Exemestane Study. After initial 2-3 years of adjuvant tamoxifen, patients were randomized between continuation up to 5 years with tamoxifen or switch to exemestane. The number of TILs was analysed for correlations with disease-free survival (DFS) and overall survival (OS). A similar analysis was performed on a cohort of 2596 Dutch patients in the TEAM trial who were randomized between the sequential scheme or exemestane monotherapy, for which follow-up was limited to the first 2.5 years in which treatments differed.

Findings In the first cohort, patients with below median number of CD8-positive TILs had a hazard ratio (HR) for DFS of 0.27 (95%CI 0.13-0.55) in favour of treatment with exemestane as compared to tamoxifen, whereas this benefit was not observed in patients with above median number of TILs (HR 1.34, 95%CI 0.71-2.50, HR for interaction 5.02, p=0.001). In the second cohort, patients with below median number of CD8-positive TILs also showed a clinical benefit of exemestane treatment on recurrence-free survival (RFS HR 0.67, 95%CI 0.45-0.99), and again not with above median number of CD8-positive TILs (HR 0.86, 95%CI 0.59-1.26, HR for interaction 1.29, p=0.36).

Interpretation This study is the first to suggest the number of CD8-positive TILs as a potential predictive marker for endocrine therapy, with a low presence of CD8-positive TILs associated to a benefit for exemestane-containing therapy. However treatment-by-marker interactions was only significant in one cohort, indicating the need for further validation.

Funding None.

Introduction

Approximately 75% of all breast cancer patients have estrogen receptor (ER)-positive tumours, and are candidates for adjuvant endocrine treatment with either an aromatase inhibitor (AI) or the selective estrogen receptor modulator (SERM) tamoxifen. Among other studies, the phase III Intergroup Exemestane Study (IES), which randomized 4724 postmenopausal patients with early stage breast cancer after 2-3 years of tamoxifen therapy between either continuing on tamoxifen or to switch to exemestane to complete 5 years of endocrine therapy, showed a significantly improved disease-free survival (DFS) for a switch to exemestane after 2-3 years of tamoxifen, compared to 5 years of tamoxifen monotherapy.[1-4] A second study, the Tamoxifen Exemestane Adjuvant Multinational (TEAM) phase 3 trial, was performed to assess the benefit of 5 year exemestane monotherapy over the switch scheme, and showed no statistical differences in survival between both groups.[5]

A recent meta-analysis in which both of the above studies are included showed that adjuvant therapy with 5 years of AI is superior to any 5 year treatment strategy with tamoxifen.[6] However, the absolute differences in recurrence and overall survival are small (between 1% and 3% on overall survival at 10 years of follow-up[6]), leaving options for biomarkers able to stratify for the benefit of either AI or tamoxifen, or predict the need for therapy extension.[7] Classic prognostic factors like TNM-stage, tumour grade, and expressional status of hormone receptors or the human epidermal growth factor receptor 2 (HER2) do not predict which adjuvant endocrine treatment is best for which patient.[5]

One of the factors that could act as a new prognostic or predictive biomarker may be derived from the immune system. The importance of the local immune system, in particular the role of tumour-infiltrating lymphocytes (TILs), on the outcome of (neo) adjuvant treatment of breast cancer has recently been validated.[8-15] Cytotoxic (CD8-positive) T-cells appear to play a major role in this phenomenon.[9,11] Most of the studies reported a clinical benefit for tumours with a higher infiltration of TILs, although this effect seems to be isolated to rapidly proliferating, ER-negative tumours.[9-14] Especially in triple negative tumours, TILs are a promising biomarker for the success of (neo)adjuvant chemotherapy.[8,15] However, no data are available which assesses the predictive value of TILs for endocrine treatment.

The aim of the current study was to determine the prognostic value of CD8-positive TILs in ER-positive breast cancer, and predictive value of CD8-positive TILs on the outcome of endocrine therapy with either tamoxifen or exemestane in two independent cohorts. For this, we evaluated the number of CD8-positive TILs in the Dutch subsets of the IES and TEAM trials, and used this for a stratified survival analysis for tumour recurrence and survival time of patients treated with either exemestane or tamoxifen.

Material and methods

Patients and tumour tissues

IES trial

In the IES trial, 4724 patients, who were treated with surgery for early breast cancer and who were disease-free after 2-3 years of adjuvant treatment with tamoxifen, were randomized between either continuing tamoxifen up to 5 years, or to switch to exemestane to complete 5 years of therapy. For the Dutch fraction of this cohort (n=236), formalin-fixed paraffin-embedded (FFPE) tumour tissue was collected and was separately converted into a tissue microarray (TMA). This TMA was created as described earlier.[16] Briefly, two 0.6mm core needle punches were obtained from the FFPE tumour blocks, and transplanted into an empty recipient block. Follow-up for disease free survival (DFS, defined as any local, regional or distant recurrence, new contralateral breast cancer or death due to any cause) and overall survival (OS) started at randomization after 2-3 years of tamoxifen treatment. For this analysis, follow-up data were used which were described earlier.[4]

TEAM trial

The TEAM trial consists of 9779 patients who were randomized for adjuvant treatment between the switch scheme (2.5 years tamoxifen followed by 2.5 years of exemestane) or 5 years of exemestane. FFPE tumour tissue was collected for the Dutch part of this trial (n=2596), and embedded in triplicate on a TMA with 0.6mm punches. Since both randomization arms were similar after the moment of switch, we censored the follow-up at 2.75 years (which was the middle between 2.5 and 3 years, the timeframe for patients in the switch group to switch to exemestane) in order to solely compare the differential effect of exemestane and tamoxifen. Beyond these 2.75 years, both treatment groups were treated with exemestane, which could interfere with the marker-by-treatment interaction. Due to the censoring at 2.75 years, only recurrence-

free survival (RFS), defined as any breast cancer recurrence or death due to breast cancer if no recurrence was reported before death, was used as a parameter of clinical outcome in this study since this censoring did not allow sufficient time to have an effect on mortality outcomes. All samples of both cohorts were handled in a coded fashion, according to national ethical guidelines ("Code for Proper Secondary Use of Human Tissue", Dutch Federation of Medical Scientific Societies).

Immunohistochemical staining

The procedures for the used immunohistochemical staining have been described before by our group in multiple different cohorts.[10, 17] In short, 4 µm sections from FFPE TMA blocks were deparaffinised in xylene and subsequently hydrated using graded alcohol washes, before endogenous peroxidase was blocked using hydrogen peroxide. Antigen retrieval was performed at 95 degrees Celsius for 10 minutes in a pH low target retrieval solution (DAKO, Glostrup, Denmark). The sections were incubated overnight at room temperature with primary antibodies against CD8 (clone 144B, Abcam, Cambridge, UK) at a predetermined optimal dilution using proper positive and negative controls. After washing, the sections were incubated with specific horseradish peroxidase-labeled Envision+ System-HRP (DAKO) for 30 minutes, before they were stained using 3,3'-diaminobenzidine (DAB) solution (DAKO). Subsequently, the slides were counterstained for 30 seconds in haematoxylin, dehydrated using inverse graded alcohol washes and xylene, and mounted in Pertex before they were dried and stored until analysis.

Evaluation of immunohistochemical staining

Slides were scanned using an automated scanner (Philips, Eindhoven, Netherlands), and obtained digital images were stored on an internal server until analysis. Each punch, of which at least 30% of the total area were tumour cells, was individually assessed for the number of CD8-positive cells in the punch by a trained investigator. Results from duplicate (IES) or triplicate (TEAM) punches were then combined in order to determine the average score per patient. The median cohort value was used as a cut-off for dichotomous analysis for infiltrating cells. Since the evaluation in the TEAM trial was intended as a proof of principle and not as a formal validation, the median value of this second cohort was used as the cut-off for this second cohort. One-third of all measurements were scored by an independent second observer, and in case of disagreement about the dichotomous classification the punch was reviewed and discussed by both observers until agreement was reached.

Statistical analysis

The study was a non-planned, retrospective, explorative project, for which all available cases were used without a predefined sample size calculation to detect a specific effect size or reach a certain level of power. ANOVA and post-hoc Bonferroni tests (corrected for multiple testing) were used to assess the mean number of CD8-positive TILs per subgroup. The kappa measurement for overall inter-observer agreement was used to assess the inter-observer variation for the dichotomized scores in one-third of all cases. Cox regression modelling was used to assess DFS and OS in the IES cohort, and RFS in the TEAM cohort, correct for possible confounders, and perform a treatment-by-marker interaction test. Missing data were included in models when they were missing in more than 10% of cases. Kaplan-Meier curves and the corresponding Log-rank tests were used to visualize these survival effects. Reverse Kaplan-Meier was used to determine the median follow-up duration. Furthermore, a post-hoc analysis was performed at which every threshold was tested to determine which cut-off point would lead to the most discriminate HR for interaction. All statistical analyses were performed using SPSS version 23 (IBM).

Role of the funding source

The original trials (IES and TEAM) were both funded by Pfizer, which had no role in this translational side study. There was no funding source for this study. The authors had full access to all the data and had the final responsibility for the decision to submit for publication.

Results

The Dutch IES-cohort consisted of 236 post-menopausal patients with early breast cancer (figure 1A). After creating the TMA, cores containing sufficient tumour tissue (>30%) were available from 190 patients. Patient and tumour characteristics are shown in table 1. The median age was 64 years (range 30–96 years). The median follow-up was 10.1 years (range 0.49–11.34 years). No significant differences in the number of CD8-positive TILs were observed between clinicopathological subgroups (Table 1). The median number of CD8-positive cells per punch was 4, which is equivalent to 14 cells/mm^2.

Table 1: The clinicopathological features of both cohorts are shown, including the mean number of CD8-positive TILs per punch for each subgroup. Statistical testing was performed using X^2, ANOVA and post-hoc bonferroni testing. Each significant association is indicated by a separate character (a, b and c). No significant differences were observed between subgroups of both cohorts.

		IES cohort patients		CD8+ TILs	TEAM cohort patients		CD8+ TILs
		n	%	Mean (n)	n	%	Mean (n)
Age	<50	3	1.6%	24	52	2.2%	9
	50-59	60	31.6%	13	713	30.4%	17*a
	60-69	60	31.6%	15	810	34.5%	16*b
	>70	67	35.3%	13	770	32.8%	12*a,b
Histological subtype	ductal	132	69.5%	13	1758	78.7%	15
	lobular	36	18.9%	15	368	16.5%	14
	other	22	11.6%	18	109	4.9%	13
	missing	-	-	-	110	-	-
Bloom & Richardson grade	grade 1	14	13.6%	11	350	16.0%	10*a,b
	grade 2	50	48.5%	11	1022	46.6%	15*b,c
	grade 3	38	36.9%	12	820	37.4%	18*a,c
	grade 4	1	1.0%	43	2	0.1%	37
	missing	87	-	-	151	-	-
Tumor size	0-3 cm	134	73.2%	15	1833	78.5%	15
	3-5 cm	38	20.8%	11	399	17.1%	14
	>5 cm	11	6.0%	9	103	4.4%	16
	missing	7	-	-	10	-	-
Nodal status	No	56	30.3%	19	714	31.3%	17
	1-3 N+	90	48.6%	11	1172	51.4%	14
	≥4 N+	39	21.1%	9	394	17.3%	15
	missing	5	-	-	65	-	-
PgR expression (>10%)	no	36	21.6%	19	509	23.0%	17
	yes	131	78.4%	13	1702	77.0%	15
	missing	23	-	-	134	-	-
HER2 expression	no	-	-	-	1991	88.4%	15*
	yes	-	-	-	261	11.6%	19*
	missing	190	-	-	93	-	-
Type of surgery	wide local excision	83	46.1%	16	1039	44.3%	17
	mastectomy	97	53.9%	12	1305	55.7%	14
	missing	10	-	-	1	-	-
Allocated treatment	Exemestane	94	49.5%	17	1187	50.6%	16
	Tamoxifen	96	50.5%	11	1158	49.4%	15

*post-hoc bonferroni test <0.05 (each association is indicated by a separate character)

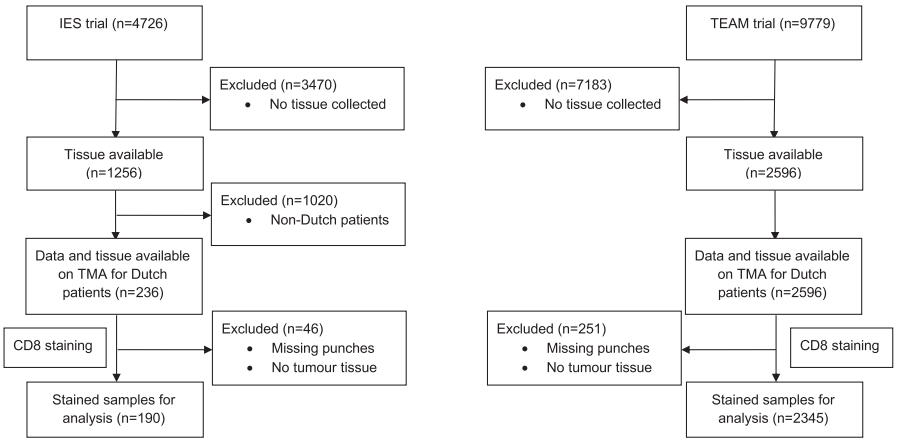


Figure 1: Flowcharts of the used cohorts for this study. The Dutch part of the Intergroup Exemestane Study (IES) (A) and the Dutch part of the international TEAM trial (B) were assessed for the number of CD8-positive TILs and its predictive value for endocrine therapy.

In the TEAM cohort, tumour tissue of 2596 patients was stained and scored for the presence of CD8-positive TILs. Sufficient cores (a minimum of 2 cores, containing at least 30% tumour tissue) were available for 2345 patients (90%). Punches showing artefacts or lack of tumour cells in the punches were excluded from analysis. The median follow-up, as determined by reverse Kaplan Meier analysis, was 2.75 years (range 0-2.75). The distribution of clinicopathological subtypes was comparable to the IES cohort (table 1). A number of significant differences in the number of CD8-positive TILs was observed between subgroups; patients above the age of 70 had a lower number of TILs compared to patients aged either 50-59 or 60-69 (Table 1). Furthermore, there was a significant association with tumour grade (more TILs with higher grade) and with HER2 expression (more TILs in HER2-positive tumours). The median number of CD8-positive TILs in this cohort was 6 per punch, which is equivalent to 20 cells/mm^2.

In the IES cohort, there was no prognostic value in the number of CD8-positive TILs for the full population for either DFS or OS (figure 2A, B). One of the aims of this study was to show the predictive value of the number of CD8-positive TILs. Therefore, we stratified the survival analysis on the number of CD8-positive TILs (table 2). It was shown that patients having a below-median number of CD8-positive TILs had a significantly better DFS when treated with exemestane after earlier tamoxifen compared to tamoxifen monotherapy (figure 3A). In 97 patients with a below-median number of CD8-positive TILs, 10 out of 45 patients on exemestane experienced a

DFS-event, whereas 31 out of 52 patients allocated to tamoxifen encountered a DFS-event. Univariate cox regression showed a hazard ratio (HR) for DFS of 0.27 (95% CI 0.13-0.55, p<0.001) in favour of exemestane treatment in these patients, with an adjusted HR (corrected for age, histological subtype, tumour size, lymph node status, tumour grade and PgR-status) of 0.35 (95% CI 0.16-0.78) (table 2). In contrast, in patients with above median numbers of CD8-positive TILs there was no significant difference in benefit of either therapy (events: 23 out of 49 on exemestane, 17 out of 44 patients on tamoxifen) with a HR of 1.34 (95% CI 0.71-2.50, p=0.36) and an adjusted HR of 1.21 (95% CI 0.58-2.51, p=0.97) (figure 2B). The HR for treatment-by-marker interaction between these groups was 5.02 (95% CI 1.93-13.02 p=0.001), showing that the difference in treatment effect between the two marker groups was statistically significant. Although underpowered due to the small cohort size and relative low numbers of events, the adjusted HR for interaction was 3.34 (95% CI 1.17-9.56, p=0.02) when corrected for age, histological subtype, tumour size, lymph node status, tumour grade and PgR-status.

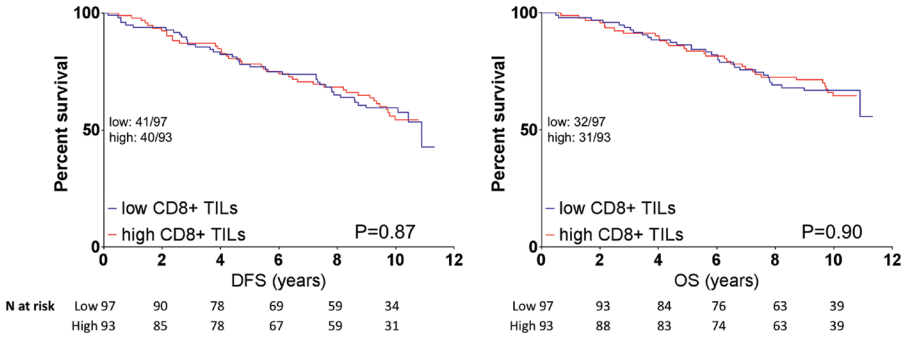


Figure 2: The general prognostic effect of CD8-positive TILs on either RFS (left) or OS (right) for all (ER-positive) patients using Kaplan-Meier survival analysis. The number of CD8-positive TILs was stratified in low (below median) and high (above median) by the median value. Event rates are provided in the graph, numbers at risk below the graph. P-values were determined using a Log-rank test.

Similar results were shown for overall survival, where a statistically significant benefit was shown for patients with a below median number of CD8-positive TILs when treated with the switch scheme. In 97 patients with a below-median number of CD8-positive TILs, 9 out of 45 patients on exemestane had died at the end of follow-up, whereas 23 out of 52 patients allocated to tamoxifen were not alive at the end of follow-up (HR 0.38, 95% CI 0.17-0.82, p=0.014; adjusted HR 0.48, 95% CI 0.19-1.18, p=0.15). In patients with an above median number of CD8-positive TILs there was no

Table 2: Results of the stratified survival analysis for recurrence free (RFS) and overall survival (OS). Cox regression was stratified based on the median number of CD8-positive TILs (below or above median). In those subgroups, we determined the hazard ratio (HR) when treated with exemestane, relative to being treated with tamoxifen. The p value for interaction indicates whether the difference in HR between the subgroups was significant. For the adjusted HR, multivariate cox regression was applied correcting for age, histological subtype, tumour size, lymph node status, tumour grade and PgR-status.

			IES cohort						TEAM cohort					
			N events/ patients	HR	95% CI	adjusted HR	95% CI	p for interaction	N events/ patients	HR	95% CI	adjusted HR	95% CI	p for interaction
RFS	Low CD8	tamoxifen	31/52	1	-	1	-	0.001** 0.02***	61/606	1	-	1	-	0.36** 0.52***
		exemestane	10/45	0.27*	0.13-0.55	0.35*	0.16-0.78		41/593	0.67*	0.45-0.99	0.71	0.47-1.07	
	High CD8	tamoxifen	17/44	1	-	1	-		56/552	1	-	1		
		exemestane	23/49	1.34	0.71-2.50	1.21	0.58-2.51		52/594	0.86	0.59-1.26	0.82	0.56-1.21	
OS	Low CD8	tamoxifen	23/52	1	-	1	-	0.04** 0.14***	-	-	-	-	-	
		exemestane	9/45	0.38*	0.17-0.82	0.48	0.19-1.18		-	-	-	-	-	
	High CD8	tamoxifen	14/44	1	-	1	-		-	-	-	-	-	
		exemestane	17/49	1.13	0.56-2.30	1.07	0.46-2.49		-	-	-	-	-	

*p<0.05

** p-value for interaction based on unadjusted analysis

*** p-value for interaction based on adjusted analysis

difference (HR 1.13, 95% CI 0.56-2.30, p=0.73; adjusted HR 1.07, 95% CI 0.46-2.49, p=0.78), with 17 out of 49 patients died on exemestane and 14 out of 44 patients died on tamoxifen (figure 3C, D). Also for overall survival, a significant treatment-by-marker interaction was observed (HR for interaction 3.01, 95% CI 1.05-8.58, p=0.04). The (underpowered) adjusted HR for interaction was 2.43 (95% CI 0.75-7.88, p=0.14).

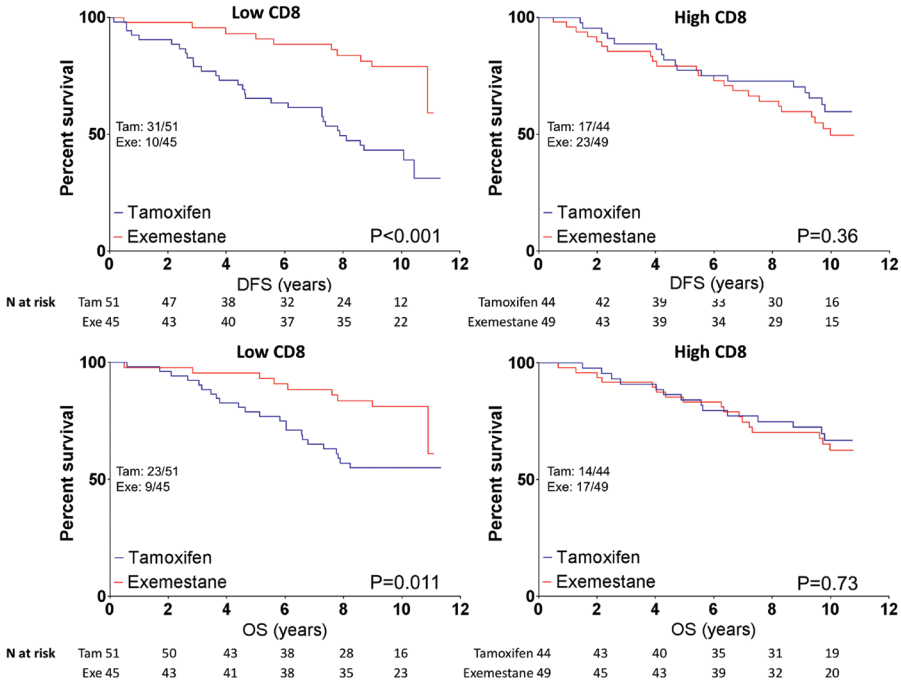
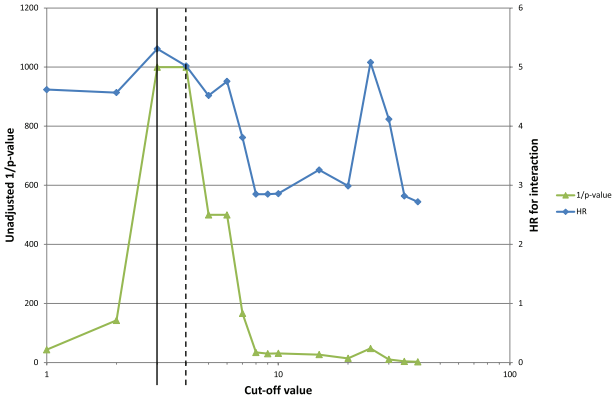


Figure 3: The predictive value of CD8-positive TILs on endocrine therapy in the IES cohort using Kaplan-Meier survival analysis. Patients with below median (low) numbers of CD8-positive TILs are shown in the left graphs (RFS above OS), patients with above median (high) numbers of CD8-positive TILs in the right graphs. Event rates are provided in the graph, numbers at risk below the graph. P-values were determined using a Log-rank test.

In a post-hoc analysis, it was established that the median value of 4 cells per punch (14 cells/mm^2) was close to the optimal threshold level of 3 cells per punch (11 cells/mm^2), which would have resulted in the highest predictive effect of CD8-positive TILs (supplemental figure 1) in the Dutch IES cohort.



Supplemental figure 1: A graph showing the HR for interaction (blue line, right y-axis) and 1/p-value (green line, left y-axis) for all cut-off values of the CD8-positive cell count, in order to determine the optimal cut-off value. The x-axis shows all possible cut-off points (cells/punch) to divide the number of CD8-positive TILs in two groups. The solid vertical line represents the optimal cut-off value of 3 cells per punch (highest HR for interaction and highest 1/p-value), whereas the dashed line represents the median value which is used as cut-off value for this study.

In order to further explore the observed interaction between the outcome of endocrine therapy and the number of CD8-positive TILs, a similar analysis was performed in the Dutch TEAM-cohort. Only the first 2.75 years of follow-up were considered for survival analysis, since after this timepoint patients in both groups received exemestane which would diminish any biological interaction.

It was established that also in this cohort, the number of TILs had no prognostic effect on recurrence either censored at 2.75 years (HR 0.91, 95% CI 0.69-1.19, p=0.47) or at full length of follow-up (HR 1.0, 95% CI 0.85-1.18 p=0.97). With regard to the predictive value, it was shown that patients with a below-median number of CD8-positive TILs, had a HR for tumour recurrence of 0.67 (95% CI 0.45-0.99, p=0.048) in favour of exemestane treatment, whereas patients with above median numbers of CD8-positive TILs had a HR of 0·86 (95% CI 0.59-1.26, p=0.44), which was similar to the findings of the first cohort (figure 4A, B). The adjusted HRs were not significant in either the CD8-low or CD8-high group (low numbers of CD8-positive TILs: 0.71, 95% CI 0.47-1.07, p=0.10; high numbers of CD8-positive TILs: 0.82, 95% CI 0.56-1.21, p=0.32). The treatment by marker interaction was not significant in this cohort (HR for interaction 1.29, 95% CI 0.75-2.22, p=0.36, adjusted HR for interaction 1.20, 95% CI 0.68-2.11, p=0.52).

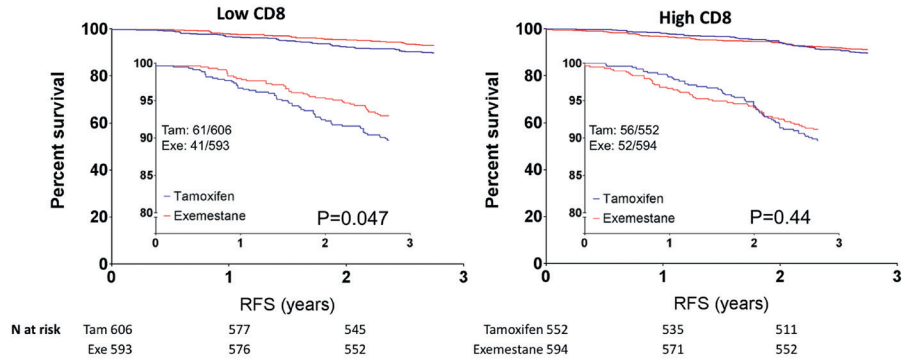


Figure 4: The predictive value of CD8-positive TILs on endocrine therapy in the TEAM cohort using Kaplan-Meier survival analysis, stratified on the median number of CD8-positive TILs. Patients with below median (low) numbers of CD8-positive TILs are shown on the left, and patients with above median (high) numbers of TILs on the right. Inserts show a more detailed graph with a range of 80-100% survival. Event rates are provided in the graph, numbers at risk below the graph. P-values were determined using a Log-rank test.

Discussion

This study is the first to investigate TILs as a predictive biomarker for the type of adjuvant endocrine therapy in postmenopausal patients with early breast cancer. In the first IES cohort, patients with a low number of CD8-positive TILs had significantly greater treatment benefit from aromatase inhibitors (AIs) than from tamoxifen, whereas the type of therapy did not make any difference in patients with high numbers of TILs. The treatment by marker interaction, comparing the clinical benefit in both subgroups, was significant despite the low number of events in this analysis, suggesting a predictive capacity of TILs for endocrine therapy. In the second TEAM cohort, it was similarly suggested that patients with low levels of CD8-positive TILs had greater treatment benefit from exemestane. However, the treatment-by-marker interaction in this cohort was not significant, indicating that the benefit of exemestane in the CD8-low group was not significantly different from the benefit in the CD8-high subgroup.

The difference in significance between both cohorts can be explained by several factors. First, the IES cohort was smaller, and thereby underpowered for definite conclusions since it is more sensitive for random variation and artefactual findings. Secondly, all patients in the IES cohort were pre-treated with 2-3 years of tamoxifen, whereas the TEAM patients were treatment-naïve at the time of randomization. This pre-treatment, and the subsequent carry-over effect known from tamoxifen, could

have influenced the differences between both cohorts. Finally, in the TEAM cohort the follow-up was censored to 2.75 years, which limited the number of events and therefore hampered the power for survival and interaction analysis. In contrast, the analysis in the IES cohort started at 2-3 years after diagnosis, and was continued up to almost 12 years post-diagnosis. This difference in follow-up periods could have influenced the comparison between both cohorts as well.

Earlier studies showed that TILs have no prognostic value in ER-positive disease.[11, 14] We confirmed these findings in both of our cohorts, showing that the number of CD8-positive TILs on itself had no prognostic value in both ER-positive cohorts. Interestingly, the suggestion that treatment with exemestane could be particularly beneficial for patients with a low number of infiltrating CD8-positive T-cells as suggested by some of our results has never been shown before in a trial-based translational study.

The mechanism behind the possible better effect of aromatase inhibitors in case of low levels of CD8 positive cells is unknown yet. Various hypothesis can be made. In accordance to our findings, one earlier study has suggested that the effect of AIs is dependent on immune suppression rather than activation.[18] In this study, Dunbier *et al.* obtained 81 paired samples before and after two weeks of neo-adjuvant anastrozole, and performed a multigene expression profile of these samples. In total, 1327 genes were differentially expressed. Although the gene expression changes varied greatly between all patients, it was observed that a higher baseline expression of pro-inflammatory genes correlated to a poor therapeutical effect of anastrozole. Upon further analysis by a pathologist, it was shown that lymphocytic infiltration correlated to a poorer therapeutical response to AIs, which was similarly observed by Tsang *et al.*[18, 19] Further on, Gao *et al.* validated these findings by showing that a high expression of genes associated with immune reaction predicted a poor response to endocrine therapy.[20]

Aromatase inhibitors might also play a role in modulating the local immune response. For example, according to the study of Generali *et al.*, aromatase inhibitors are capable of lowering the number of tumour-infiltrating regulatory T-cells, and thereby may improve treatment outcome.[21] Similar results were shown by Chan *et al.*, who studied the ratio of cytotoxic T-cells and regulatory T-cells during neoadjuvant endocrine treatment and observed a significant increase of this ratio in responders, as opposed to non-responders.[22] Moreover aromatase inhibitors have been shown to enhance cytokine excretion and the severity of experimental polyarthritis in murine models, indicating

an activation of the immune system.[23] Furthermore, auto-immune conditions have been suggested as a contributing factor to often reported arthralgia.[24] Based on these abovementioned findings, it could be hypothesized that aromatase inhibitors exert part of their function by activating both the systemic and the local immune response. Therefore, patients with a weaker local immune response at baseline will benefit more from AIs, since the immunomodulation will yield more effect in those patients compared to patients who already have a strong local immune response.

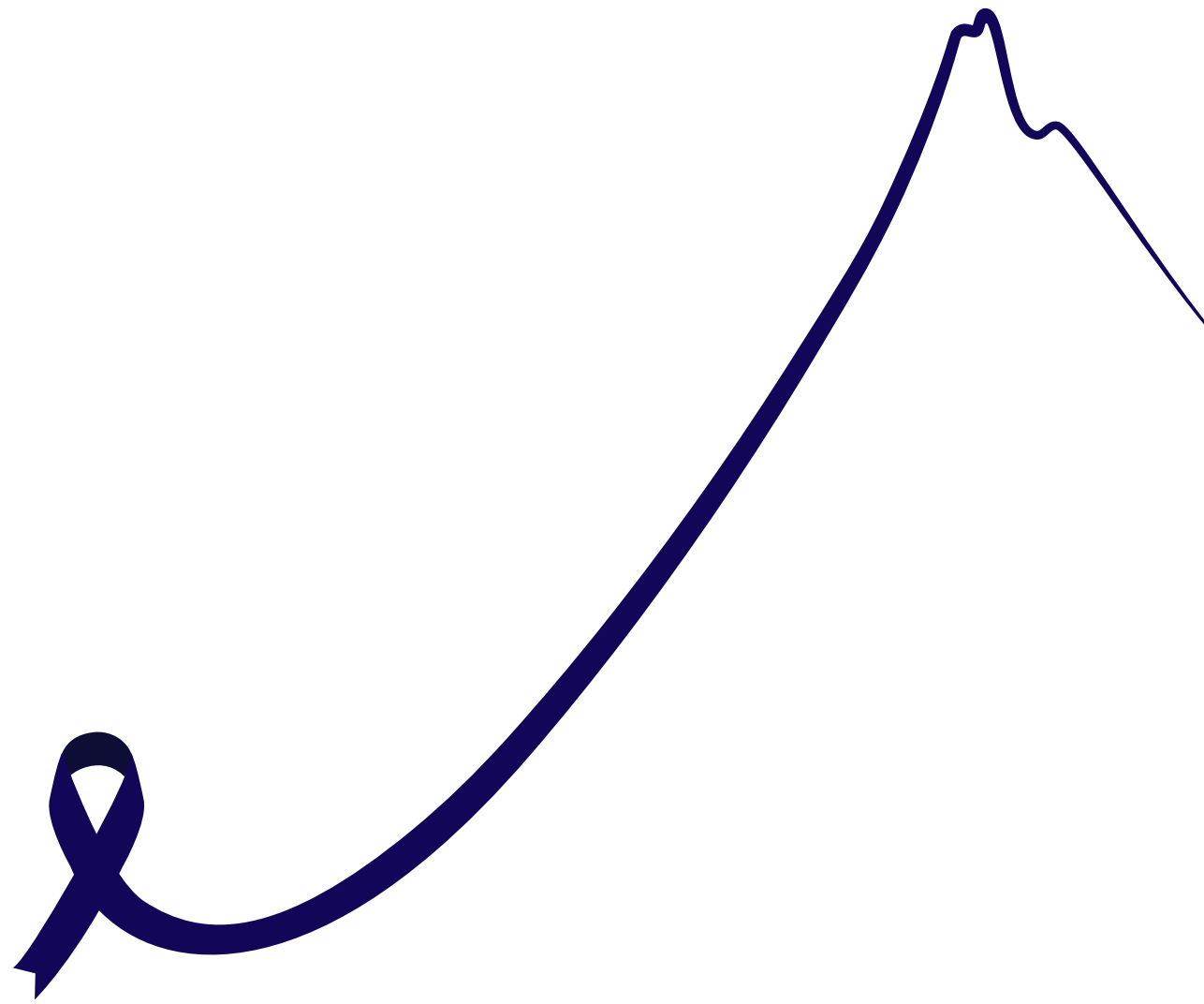
Another theory for explaining the possible differential effect of AIs and tamoxifen between TIL-rich and TIL-poor tumours, is that the number of infiltrating CD8+ TILs is a proxy variable for another tumour characteristic, which might be the mutational load. Earlier, it was established that the mutational load in the tumour, and therefore the number of neo-epitopes, is associated with the local immune response.[25] Furthermore, it has been shown that more aggressive Luminal B-type tumours, which are generally considered less responsive to endocrine therapy, have a higher mutational load compared to the more responsive Luminal A subtype.[26, 27] Hypothetically, tumours with a lower mutational load might be more dependent on ER-pathway signalling since they are less likely to acquire activating mutations in other oncogenes, whereas tumours with a higher mutational load have activated other growth stimulating pathways and are therefore less dependent on ER-signalling for their survival. These results suggest that AIs would be the most optimal strategy for strongly ER-dependent (lower mutational load) tumours, whereas tamoxifen and AIs are equally good for less ER-dependent tumours.

In summary, the current study provides the first suggestion that the number of CD8-positive TILs could be used as a predictive marker in the endocrine treatment of breast cancer. Upon further validation in a trial with a similar design as IES in which tamoxifen monotherapy is compared to an AI-containing regime, patients with low numbers of CD8-positive TILs could have more benefit from AIs than from tamoxifen, whereas patients with a strong infiltration of CD8-positive TILs have a similar outcome on both treatment strategies. Future studies will be directed towards validation of these findings for other aromatase inhibitors, to show whether the results observed for exemestane can be extrapolated to letrozole or anastrozole as well. Our findings might contribute to a more optimized treatment of hormone-receptor-positive breast cancer using the local immune system as a predictive biomarker for adjuvant endocrine therapy.

References

1. Coombes RC, Kilburn LS, Snowdon CF et al. Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (intergroup exemestane study): a randomised controlled trial. Lancet 2007;369:559-570.
2. Bliss JM, Kilburn LS, Coleman RE et al. Disease-related outcomes with long-term follow-up: an updated analysis of the intergroup exemestane study. J Clin Oncol 2012;30:709-717.
3. Coombes RC, Hall E, Gibson LJ et al. A Randomized Trial of Exemestane after Two to Three Years of Tamoxifen Therapy in Postmenopausal Women with Primary Breast Cancer. N Engl J Med 2004;350(11):1081-1092.
4. Morden JP, Alvarez I, Bertelli G et al. Long-Term Follow-Up of the Intergroup Exemestane Study. JCO 2017;JCO.
5. van de Velde CJ, Rea D, Seynaeve C et al. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. The Lancet 2011;377(9762):321-331.
6. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
7. Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.
8. Loi S, Michiels S, Salgado R et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Annals of Oncology 2014;25(8):1544-1550.
9. Ali HR, Provenzano E, Dawson SJ et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. Annals of Oncology 2014;25(8):1536-1543.
10. de Kruijf EM, Engels CC, van de Water W et al. Tumor immune subtypes distinguish tumor subclasses with clinical implications in breast cancer patients. Breast Cancer Res Treat 2013;142:355-364.
11. Mahmoud SMA, Paish EC, Powe DG et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. JCO 2011;29(15):1949-1955.
12. West NR, Milne K, Truong PT et al. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res 2011;13(6):R126.
13. West NR, Kost SE, Martin SD et al. Tumour-infiltrating FOXP3(+) lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. Br J Cancer 2013;108(1):155-162.
14. Baker K, Lachapelle J, Zlobec I et al. Prognostic significance of CD8+ T lymphocytes in breast cancer depends upon both oestrogen receptor status and histological grade. Histopathology 2011;58(7):1107-1116.
15. Adams S, Gray RJ, Demaria S et al. Prognostic Value of Tumor-Infiltrating Lymphocytes in Triple-Negative Breast Cancers From Two Phase III Randomized Adjuvant Breast Cancer Trials: ECOG 2197 and ECOG 1199. JCO 2014.
16. Camp RL, Charette LA, Rimm DL. Validation of Tissue Microarray Technology in Breast Carcinoma. Lab Invest 2000;80(12):1943-1949.
17. Engels CC, Charehbili A, van de Velde CJH et al. The prognostic and predictive value of Tregs and tumor immune subtypes in postmenopausal, hormone receptor-positive breast cancer patients treated with adjuvant endocrine therapy: a Dutch TEAM study analysis. Breast Cancer Res Treat 2015;149(3):587-596.
18. Dunbier AK, Ghazoui Z, Anderson H et al. Molecular Profiling of Aromatase Inhibitor-Treated Postmenopausal Breast Tumors Identifies Immune-Related Correlates of Resistance. Clinical Cancer Research 2013;19(10):2775-2786.
19. Tsang J, Hui SW, Ni YB et al. Lymphocytic infiltrate is associated with favorable biomarkers profile in HER2-overexpressing breast cancers and adverse biomarker profile in ER-positive breast cancers. Breast Cancer Res Treat 2014;143(1):1-9.
20. Gao Q, Patani N, Dunbier AK et al. Effect of Aromatase Inhibition on Functional Gene Modules in Estrogen Receptor-Positive Breast Cancer and Their Relationship with Antiproliferative Response. Clinical Cancer Research 2014;20(9):2485-2494.
21. Generali D, Bates G, Berruti A et al. Immunomodulation of FOXP3+ regulatory T cells by the aromatase inhibitor letrozole in breast cancer patients. Clin Cancer Res 2009;15:1046-1051.

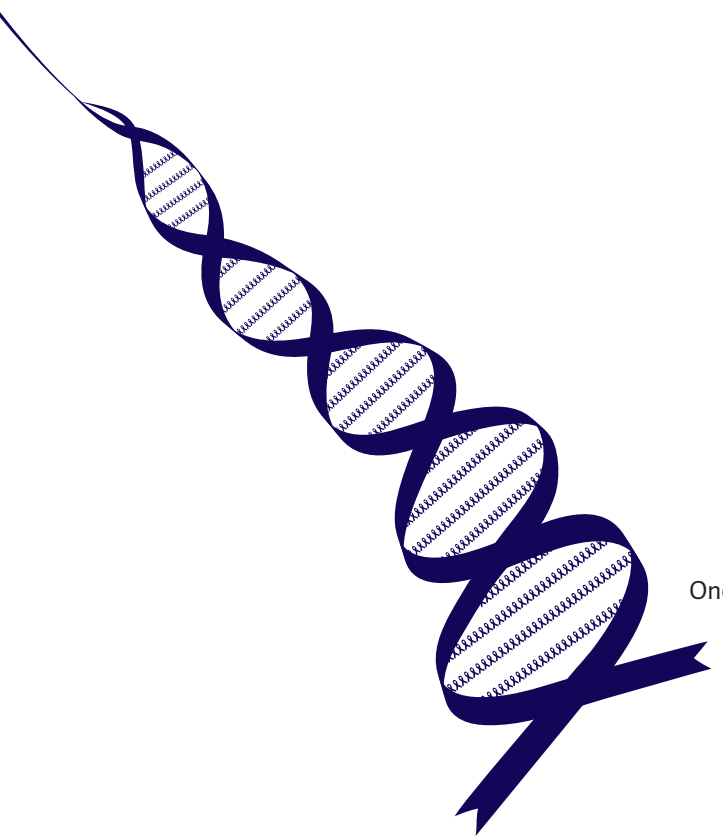
22. Chan MS, Wang L, Felizola SJ et al. Changes of tumor infiltrating lymphocyte subtypes before and after neoadjuvant endocrine therapy in estrogen receptor-positive breast cancer patients--an immunohistochemical study of Cd8+ and Foxp3+ using double immunostaining with correlation to the pathobiological response of the patients. Int J Biol Markers 2012;27(4):e295-e304.
23. Jingxuan W, Qingyuan Z, Shi J et al. Immoderate inhibition of estrogen by anastrozole enhances the severity of experimental polyarthritis. Experimental Gerontology 2009;44(6-7):398-405.
24. Wendling D, Letho-Gyselinck H+, Guillot X et al. Arthralgia and aromatase inhibitors. Joint Bone Spine 2014;81(2):187-188.
25. McGranahan N, Furness AJ, Rosenthal R et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351(6280):1463-1469.
26. Creighton CJ. The molecular profile of luminal B breast cancer. Biologics 2012;6:289-297.
27. Budczies J, Bockmayr M, Denkert C et al. Classical pathology and mutational load of breast cancer: integration of two worlds. J Path: Clin Res 2015;1(4):225-238.



Chapter 8

Combined evaluation of the FAS cell surface death receptor and CD8+ tumor infiltrating lymphocytes as a prognostic biomarker in breast cancer

E.J. Blok
J. van den Bulk
N.G. Dekker-Ensink
R. Derr
C. Kanters
E. Bastiaannet
J.R. Kroep
C.J.H. van de Velde
P.J.K. Kuppen



Oncotarget, 2017, 28;8(9): 15610-15620

Abstract

Multiple studies showed the prognostic capacities of tumor-infiltrating lymphocytes (TILs) in triple-negative breast cancer (TNBC), but not in other subtypes. We evaluated tumor expression of FAS, a key receptor in T-cell mediated apoptosis, as possible explanation for this differential prognostic value of TILs. Furthermore, we evaluated the prognostic relevance of FAS, both as an independent biomarker and in relation to CD8-positive T-cell presence. The study cohort consisted of 667 breast cancer patients treated in the LUMC between 1997 and 2009. FAS expression was determined using immunohistochemistry and the percentage of FAS-positive tumor cells was quantified. Furthermore, the number of CD8-positive infiltrating cells was determined, and its prognostic relevance was associated to FAS-expression using stratified survival analysis. In TNBC, FAS was averagely expressed in 49% of tumor cells, whereas ER-positive subtypes showed an average Fas expression of 16-20%. In the entire cohort, FAS was identified as significant prognostic marker for recurrence (adjusted HR 0.53, 95% CI 0.36-0.77) and borderline significant marker for overall survival (adjusted HR 0.72, 95% CI 0.52-1.01). Upon stratification for FAS-expression, CD8+ TILs were only prognostic at high levels (above median) of FAS expression in ER-negative disease. In summary, FAS was identified as an independent prognostic marker for recurrence free survival in breast cancer, with large variation in expression by receptor subtypes. Interestingly, the prognostic effect of CD8+ TILs in ER-negative disease was only valid for tumors with a high FAS expression.

Introduction

For decades the local immune response, among others represented by quantification of tumor-infiltrating lymphocytes (TILs), has been studied in breast cancer.[1-3] Although most studies observed a prognostic value of TILs, these studies have not resulted in any change in clinical practice. Studies have shown TILs to have strong prognostic impact in ER-negative, HER2-negative and triple negative breast cancer (TNBC), but not in ER-positive tumors.[4-14] In a recent meta-analysis by Ibrahim *et al* combining the results of 8 studies, a 30% reduction in disease recurrences and a 22% decrease in distant recurrences was shown for triple-negative patients having high amounts of TILs.[14] Furthermore, a hazard ratio 0.66 (95% CI 0.53-0.83) for overall survival was shown for these patients, providing robust evidence for the prognostic value of TILs.

It is known that although TILs might be present in the tumor, their functioning might be hampered [3]. One of the most studied factors involved is classical HLA class I, which was shown to be downregulated in breast cancer and other malignancies.[15, 16] Another protein on tumor cells that determines function of T cells is Fas cell surface death receptor, abbreviated as FAS. FAS is broadly expressed on most normal tissue, and is a crucial link between T-cell mediated immunity and induction of apoptosis.[17, 18] When a cytotoxic T-cell binds to a target cell, FAS-ligand (FASL) is upregulated by the T-cell. FASL subsequently binds to the target cell-expressed FAS, thereby initiating the activation of a caspase cascade leading to apoptosis of the target cell. Together with perforin-induced apoptosis, these are the two main mechanisms by which a cytotoxic T-cell can induce apoptosis.[19, 20] It could be hypothesized that downregulation of FAS is a mechanism of tumor immune evasion, since this disables a crucial step in T-cell mediated immunity. Therefore, tumor expression of FAS could act as a clinical prognostic marker in breast cancer.

Hypothetically, the expression of FASL by tumor cells could lead to induction of apoptosis in the cytotoxic T-cells which could be a second method of FAS-FASL-mediated immune evasion. A number of studies have been performed evaluating the prognostic relevance of FAS and FASL in breast cancer, focusing mainly on the FASL/FAS ratio.[21-23] These studies indeed reported that a higher tumor expression of FASL and/or a lower expression of FAS, resulting in an increased FASL/FAS ratio, associated with a worse disease free and overall survival.[21] Other studies reported that this was mainly due to an increase in FAS-expression, whereas FASL did not influence

outcome.[23] Furthermore, the theory of immune evasion by upregulation of FASL in the tumor has never been shown *in vivo*.[24] Therefore, it is expected that most effects seen for the FASL/FAS ratio in tumors are attributed to a downregulation of FAS.

Although TILs have shown to be of prognostic relevance, it is highly unlikely that the TILs in the primary tumor will determine survival outcome. Most likely the amount of TILs in the primary tumor is a proxy variable for a yet undefined tumor characteristic, making the tumor more or less susceptible for an immune response. This process could lead to an aberrant pattern of metastasizing, or an effect on growth speed of the metastasis. When FAS is differentially expressed among different tumor subtypes, it could be hypothesized that FAS is a key explanatory factor for the fact that TILs are prognostic in one subgroup, but not in other subgroups. Furthermore, combining recent evidence regarding TILs in TNBC with the earlier evidence on FAS expression, we suggest that FAS is a clinical prognostic in breast cancer as an independent alternative for TILs.

Therefore, three main aims of this study are identified: To evaluate the expression of FAS among different tumor subtypes in order to explain variances in the prognostic value of TILs. The second aim is to evaluate the expression of FAS as a prognostic marker in breast cancer, both in general and in selected subtypes. Finally, the third aim of this study was to evaluate the prognostic value of CD8 in the presence or absence of FAS-expression, since we hypothesize that CD8-positive T-cells will only be prognostic in the presence of tumor FAS expression.

Results

Baseline characteristics

667 patients were included in this observational cohort of patients treated in the LUMC (Table 1). Most tumors were categorized as ductal carcinomas (80,8%); 10,2% were determined to be lobular carcinoma. Approximately 75% of the tumors showed ER positivity, 55% PR positivity and 25% HER2 positivity. For HER2 expression, nearly 50% of the records was missing. Missing of these data was strongly correlated to the year of diagnosis. Before 2003, 89% of HER2 scores was missing (322 of 360 patients), whereas from 2003 onwards it was only missing in 5% of patients (14 of 307 patients). The percentage of triple negative tumors was 16%, whereas ER+PR+HER2- was the most prevalent subtype with 42%. The majority of tumors were small and early stage

Table 1 – Baseline overview of the clinicopathological parameters of the cohort. Both FAS-expression and presence of CD8+ tumor infiltrating lymphocytes (TILs) are shown, stratified according to standard clinicopathological parameters. Percentages are excluding missing variables. *column proportion test p-value <0.05

		Cohort description		Fas expression by median N=640 (27 excluded)				Fas expression	CD8 in tumor by median N=625 (42 excluded)				CD8+ TILs
		Total N=667		Low (< median)		High (> median)		Mean (%)	Low (< median)		High (> median)		Mean (n)
		N	%	N	%	N	%		N	%	N	%	
Age	<40	55	8,2%	17	30,9%	38	**69,1%***	34	13	24,1%	41	**75,9%***	42
	40-49	153	22,9%	67	46,2%	78	53,8%	23	58	40,8%	84	**59,2%***	27
	50-59	210	31,5%	96	47,3%	107	52,7%	22	105	53,0%	93	47,0%	20
	60-69	127	19,0%	70	**56,5%***	54	43,5%	19	72	**61,0%**	46	39,0%	17
	>70	122	18,3%	56	49,6%	57	50,4%	21	65	57,5%	48	42,5%	19
Histological subtype	ductal	539	80,8%	248	48,1%	268	51,9%	23	255	51,0%	245	49,0%	23
	lobular	68	10,2%	32	47,8%	35	52,2%	18	30	45,5%	36	54,5%	22
	other	50	9,0%	26	45,6%	31	54,4%	23	28	47,5%	31	52,5%	25
Bloom & Richardson grade	grade 1	108	18,2%	47	46,5%	54	53,5%	20	61	**62,2%**	37	37,8%	14
	grade 2	275	46,5%	136	52,3%	124	47,7%	19	137	54,6%	114	45,4%	16
	grade 3	209	35,3%	89	43,4%	116	56,6%	29	82	40,4%	121	**59,6%***	36
	missing	75	-	-	-	-	-	-	-	-	-	-	-
ER expression (>10%)	no	140	23,5%	48	35,3%	88	**64,7%***	37	51	38,3%	82	**61,7%***	34
	yes	456	76,5%	218	49,9%	219	50,1%	19	225	53,3%	197	46,7%	20
	missing	71	-	-	-	-	-	-	-	-	-	-	-
PgR expression (>10%)	no	261	45,2%	104	41,4%	147	**58,6%***	29	104	42,6%	140	**57,4%***	30
	yes	316	54,8%	153	50,5%	150	49,5%	19	165	56,3%	128	43,7%	18
	missing	90	-	-	-	-	-	-	-	-	-	-	-
HER2 expression	no	247	74,6%	106	44,4%	133	55,6%	23	124	53,0%	110	47,0%	24
	yes	84	25,4%	45	55,6%	36	44,4%	21	40	51,3%	38	48,7%	21
	missing	336	-	-	-	-	-	-	-	-	-	-	-

Table 1 – continued

		Cohort description		Fas expression by median N=640 (27 excluded)				Fas expression	CD8 in tumor by median N=625 (42 excluded)				CD8+ TILs
		Total N=667		Low (< median)		High (> median)		Mean (%)	Low (< median)		High (> median)		Mean (n)
		N	%	N	%	N	%		N	%	N	%	
Receptor subtype	ER-PR-HER2-	52	16,0%	8	17,4%	38	**82,6%***	49	16	34,8%	30	**65,2%***	50
	ER-PR-HER2+	31	9,5%	15	48,4%	16	51,6%	29	11	37,9%	18	62,1%	19
	ER+PR-HER2-	47	14,4%	25	53,2%	22	46,8%	16	28	60,9%	18	39,1%	15
	ER+PR-HER2+	25	7,7%	13	52,0%	12	48,0%	16	13	52,0%	12	48,0%	17
	ER-PR+HER2-	9	2,8%	4	57,1%	3	42,9%	20	3	50,0%	3	50,0%	16
	ER+PR+HER2-	137	42,0%	68	49,6%	69	50,4%	18	76	56,7%	58	43,3%	18
	ER+PR+HER2+	25	7,7%	17	**68%***	8	32,0%	16	16	66,7%	8	33,3%	27
	missing	341	-	-	-	-	-	-	-	-	-	-	-
Tumor stage based on pT, pN and p/cM	IA	248	39,3%	109	46,4%	126	53,6%	21	141	**60,5%***	92	39,5%	17
	IB	4	0,6%	3	75,0%	1	25,0%	9	1	33,3%	2	66,7%	8
	IIA	184	29,2%	84	46,7%	96	53,3%	25	77	44,3%	97	55,7%	30
	IIB	133	21,0%	60	48,4%	64	51,6%	22	52	41,9%	72	**58,1%***	23
	IIIA	23	3,6%	11	50,0%	11	50,0%	27	7	33,3%	14	66,7%	35
	IIIB	5	0,8%	4	80,0%	1	20,0%	4	1	20,0%	4	80,0%	21
	IIIC	30	4,7%	15	50,0%	15	50,0%	21	16	55,2%	13	44,8%	24
	IV	4	0,6%	4	100,0%	0	0,0%	5	2	50,0%	2	50,0%	18
	missing	35	-	-	-	-	-	-	-	-	-	-	-
Neoadjuvant systemic therapy	CT	32	4,8%	12	40,0%	18	60,0%	26	16	59,3%	11	40,7%	19
	HT	22	3,3%	12	54,5%	10	45,5%	19	11	50,0%	11	50,0%	16
	CT + HT	1	0,1%	1	100,0%	0	0,0%	3	0	0,0%	1	100,0%	45
	none	612	91,8%	281	47,9%	306	52,1%	23	286	49,7%	289	50,3%	23

(stage II or lower), only approximately 10% was stage III or IV. Most patients (91,8%) did not receive neoadjuvant systemic therapy. These percentages show that the cohort is representative for the general breast cancer population.

FAS-expression

From the 667 patients included in this observational cohort, immunohistochemical staining for FAS expression was successful for 640 patients. 27 patients were excluded due to a lack of tumor tissue on the TMA, either as an artefact or because only non-tumorous tissue was included on the TMA (Figure 1). In the remaining 640 patients, FAS expression was observed ranging from 0% to 100% of the tumor cells, with a median expression of 13.3% (Figure 2). The correlations with baseline characteristics are shown in Table 1. A small difference in FAS expression was shown for age, in which younger (<40 years) patients showed a higher expression of FAS, whereas patients between 60 and 69 showed a slightly lower FAS-expression (column proportion test p-value <0.05). No associations were found for histological subtype or tumor stage. It was observed that grade 3 tumors had a significantly higher FAS expression compared to grade 1 and 2. ER-negative tumors showed almost a doubling of the average expression of FAS compared to ER-positive tumors (37% vs 19% FAS-positive tumor cells per sample, p<0.05)). For HER2, limited data were available (n=320), showing no statistical differences. Combining ER, PR and HER2, it was shown that triple negative tumors showed significantly higher FAS-expression (average of 49% positive tumor cells) compared to the other subtypes (Bonferroni multiple comparisons test p-values <0.001), especially ER-positive subtypes (FAS expression ranging from 16% to 18% positive tumor cells) (Figure 3). Pre-treatment with either neo-adjuvant chemotherapy or endocrine therapy was not associated with different FAS-expression patterns.

In the provisional TCGA dataset, levels of FAS expression were compared to the expression of ESR1, the gene encoding for ER. We observed a Pearson correlation of -0.35, meaning that a higher FAS expression is correlated to a lower expression of ER. This is in accordance with our findings that in ER-negative tumors, there is a higher FAS expression. To supplement these findings, we analyzed the TCGA dataset as published in Nature in 2012, for which more clinical data are available.[25] In this cohort, we observed that of the 14 patients who have an upregulation of FAS at transcriptional mRNA level , 13 of them are ER-negative and for one patient ER-staining was not performed. In contrast, of the 17 patients with a downregulation of FAS, 13 were ER-positive. This further validated our finding, that high levels of FAS are associated to low levels of ER-expression, both at transcriptional and protein level.

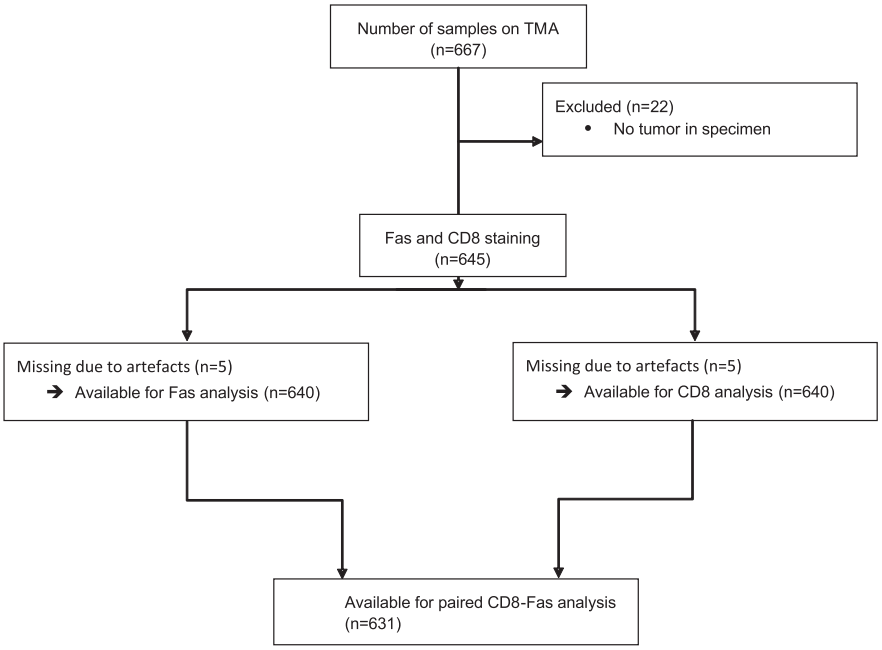


Figure 1 – Consort diagram of the included patients, which were present on the TMA, for analysis. The causes for missing samples were a lack of tumor in the punches or artefacts like folded or missing parts of the punches.

FAS-expression as clinical prognostic marker

To evaluate the clinical prognostic value of FAS-expression, Kaplan Meier curves were plotted for the general study population (Figure 4A,B). It was shown that a high FAS expression correlated with a longer recurrence free and overall survival time (log-rank p-values of 0.009 and 0.02 respectively) in the entire cohort. In a univariate cox-regression analysis, a hazard ratio of 0.65 (95% CI 0.47-0.90, p=0.01) and 0.72 (95% CI 0.55-0.95, p=0.02) was seen for RFS and OS respectively. In a multivariate cox regression analysis, corrected for age, histological subtype, tumor grade, tumor stage, ER-expression, year of diagnosis, neo-adjuvant treatment, adjuvant chemotherapy, and adjuvant endocrine therapy, an adjusted HR of 0.53 (95% CI 0.36-0.77, p=0.001) was seen for RFS. For OS, an adjusted HR of 0.72 was observed, with a borderline significance (95% CI 0.52-1.01, p=0.055).

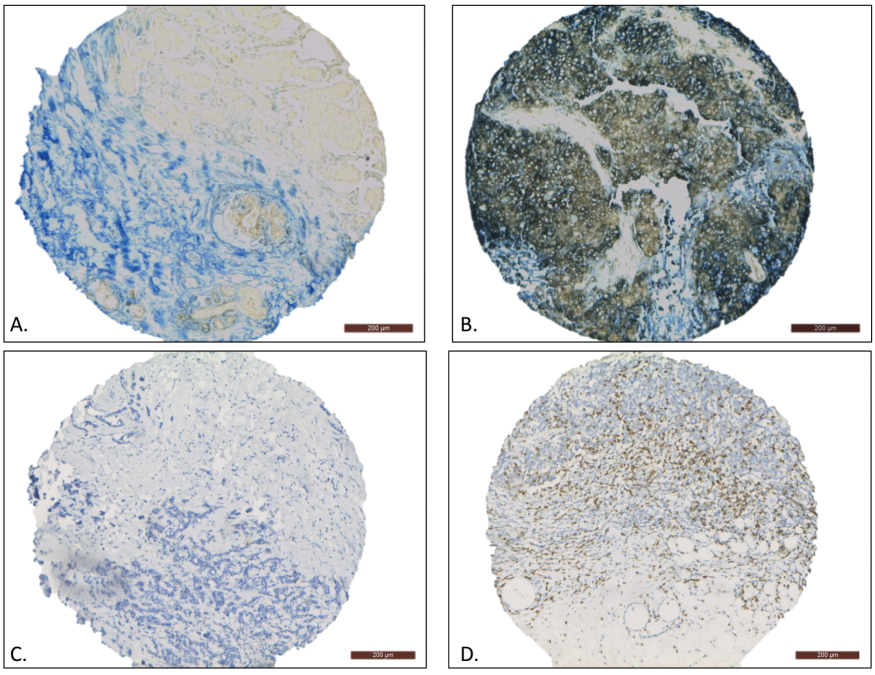


Figure 2 – Representative examples of immunohistochemical staining for FAS expression and CD8-positive TILs (10x magnification). The FAS-negative sample only contains some FAS-positive infiltrating lymphocytes (A), whereas the FAS-positive sample shows homogenous membranous FAS expression in the tumor cells (B). The CD8-low sample showed no infiltration of CD8-positive TILs (C), whereas the CD8-high sample shows large numbers of CD8-positive TILs (D).

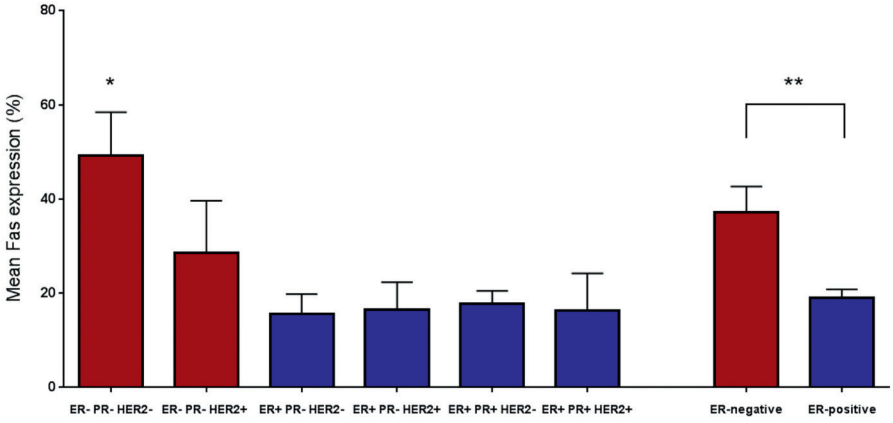


Figure 3 – The average percentage of FAS expressing tumor cells, as determined by immunohistochemical staining, according to molecular subtypes. *Significantly different from all other individual subgroups using Bonferroni's multiple comparisons test (all adjusted p-values <0.001). ** Unpaired t-test p-value <0.001

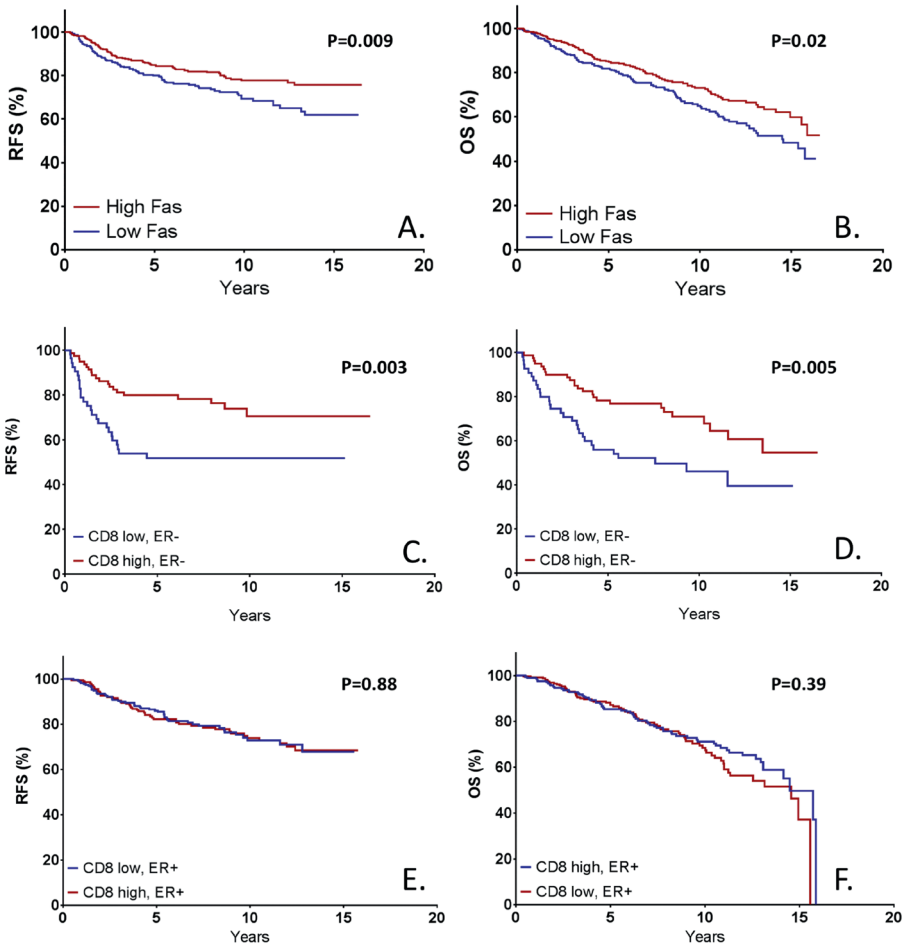
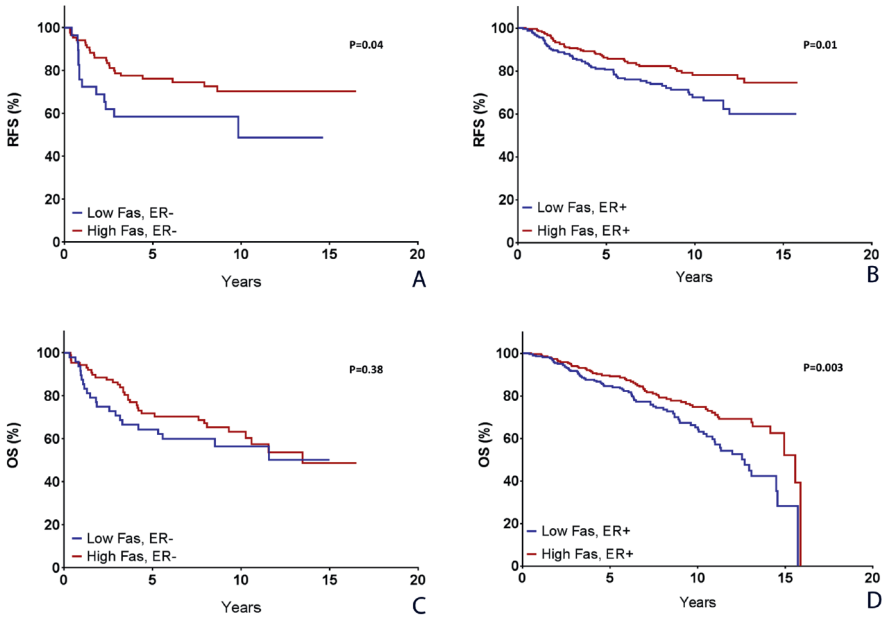


Figure 4 – Kaplan Meier survival estimates based on immunohistochemical FAS expression both for recurrence free survival (A) and overall survival (B). Furthermore, the prognostic effect of CD8-positive TILs is shown in both ER-negative (C. RFS, D. OS) and ER-positive tumors (E. RFS, F. OS). P-values represent log-rank survival test.

Upon stratification on ER-expression, it was shown that both in ER-negative (HR 0.48, 95% CI 0.27-0.86, p=0.01) and ER-positive (HR 0.65, 95% CI 0.43-0.97, p=0.04) tumors, FAS expression was prognostic for RFS (Supplemental fig S2A,B). In multivariate analysis, a HR of 0.36 (95% CI 0.17-0.76, p=0.01) was shown in ER-negative tumors, whereas a HR of 0.58 (95% CI 0.37-0.90, p=0.02) was shown for ER-positive tumors. For OS, no statistical significant differences regarding level of FAS expression were shown for ER-negative tumors in univariate (HR 0.78, 95% CI 0.45-1.35, p=0.38) or multivariate (HR 0.65, 95% CI 0.33-1.30, p=0.23) modelling (Supplemental fig S2C).

In ER-positive tumors, a strong benefit of FAS-expression was shown for OS (HR 0.59, 95% CI0.42-0.83, p=0.003), but this failed to show in multivariate analysis (HR 0.76, 95% CI 0.52-1.12, p=0.16) (Supplemental fig S2D). No significant interaction between ER-status and FAS-expression was observed for either RFS or OS (HRs 0.70 (p=0.32) and 1.15 (p=0.66) respectively), meaning that the effect of FAS expression on survival is not significantly different between ER-negative and ER-positive patients.

In summary, an above median level of FAS expression was a statistically significant independent prognostic marker for RFS, and a borderline significant prognostic marker for OS. Both effects were conserved in ER-negative and ER-positive tumors.



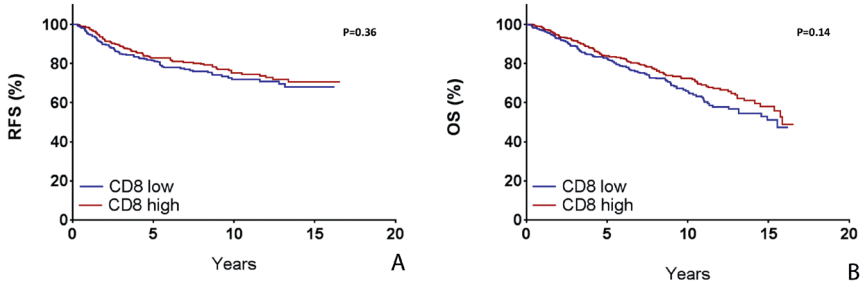
Supplementary Figure 2: The prognostic effect of FAS expression, stratified on ER-status. The effect is conserved in both ER-negative A, C. and ER-positive tumors B, D. for both RFS and OS. P-values represent log-rank survival test.

CD8+ lymphocyte infiltration

For the evaluation of CD8+ tumor infiltrating lymphocytes (TILs), TILs were counted in both the tumor and the directly adjacent stromal tissue, only when the punch contained tumor tissue. Therefore, 27 cases were excluded due to artefacts (missing punches) or a lack of tumor tissue in the sample. For the remaining 640 patients, the number of CD8+ TILs ranged from 0 to 369, with a median value of 28 per punch (1mm).

The distribution of CD8+ TIL infiltration over basic clinicopathological subgroups is shown in Table 1. Young patients (<40 and 40-49) showed an increased amount of infiltration compared to other age categories. Furthermore, ER-negative and grade 3 tumors showed increased rates of infiltration.

Over the whole cohort, CD8+ TILs showed no correlation with RFS (HR 1.16, 95% CI 0.84-1.60, p=0.36) (Supplementary Fig. S1A). However, when corrected for age, histological subtype, tumor grade, tumor stage, ER-expression, year of diagnosis, neo-adjuvant treatment, adjuvant chemotherapy, and adjuvant endocrine therapy, an adjusted HR of 0.55 (95% CI 0.37-0.81, p=0.003) was observed for RFS for patients with an above-median level of CD8+ TILs. Similar effects were shown for OS (HR 1.23, 95% CI 0.94-1.62, p=0.14; adjusted HR 0.78, 95% CI 0.56-1.10, p=0.16), although not statistically significant (Supplementary Fig. S1B).



Supplementary Figure 1: Kaplan Meier survival analysis, comparing groups with either low or high presence of CD8-positive TILs in the general study population for either RFS A. or OS B. P-values represent log-rank survival test.

In earlier studies, it was observed that CD8+ TILs were only prognostic in ER-negative or triple negative breast cancer.[4] Upon stratification on ER-expression, similar results were observed in this cohort (Figure 4C-F). In ER-positive patients, high levels of CD8+ TILs were associated with a HR for recurrence of 1.03 (95% CI 0.69-1.54, p=0.88), whereas in ER-negative disease a HR of 0.42 (95% CI 0.23-0.76, p=0.004) was observed (HR for interaction 2.64, p=0.007). A similar pattern was observed for OS (ER+ HR 0.86, 95% CI 0.61-1.21, p=0.39; ER- HR 0.48, 95% CI 0.28-0.81, p=0.007).

Combined FAS-CD8 analysis

In order to determine the hypothesized pivotal effect of FAS expression on the function, and therefore prognostic effect of CD8+ TILs, a survival analysis was performed on the presence of CD8+ lymphocytes, stratified on FAS expression. In the

complete cohort, there was no difference between the prognostic relevance of CD8 TILs for either high or low expression of FAS on RFS (HR for interaction 1.51, p=0.24) or OS (HR for interaction 1.24, p=0.45).

Upon stratification on ER expression, a similar analysis was performed. In ER-negative disease, it was shown that CD8 was prognostic for RFS in the presence of high FAS expression (HR 0.42, 95% CI 0.19-0.96, p=0.04), but not with low FAS expression (HR 0.54, 95% CI 0.23-1.29, p=0.17), with a HR for interaction of 0.80, (p=0.71) (Figure 5A,B). In ER-positive disease, CD8 was reversely prognostic for RFS with at high levels of FAS expression (HR 2.01, 95% CI 1.01-4.04, p=0.05) and not prognostic at low FAS expression (HR 0.80, 95% CI0.46-1.39, p=0.43), with a HR for interaction of 2.42, p=0.05 (Figure 5C,D). For OS, a similar pattern was observed (ER-negative, high FAS: HR 0.49 (95% CI 0.25-0.97, p=0.04), ER-negative, low FAS: HR 0.51 (95% CI 0.20-1.26, p=0.14) (Supplementary Fig. S3A,B); ER-positive, high FAS: HR 1.22 (95% CI 0.70-2.12, p=0.48), ER-positive, low FAS: HR 0.78 (95% CI 0.48-1.25, p=0.30) (Supplementary Fig. S3C,D). In summary, CD8+ TILs were only prognostic for both RFS and OS in ER-negative tumors at high levels of FAS, but not in tumors with low expression of FAS or in ER+ tumors.

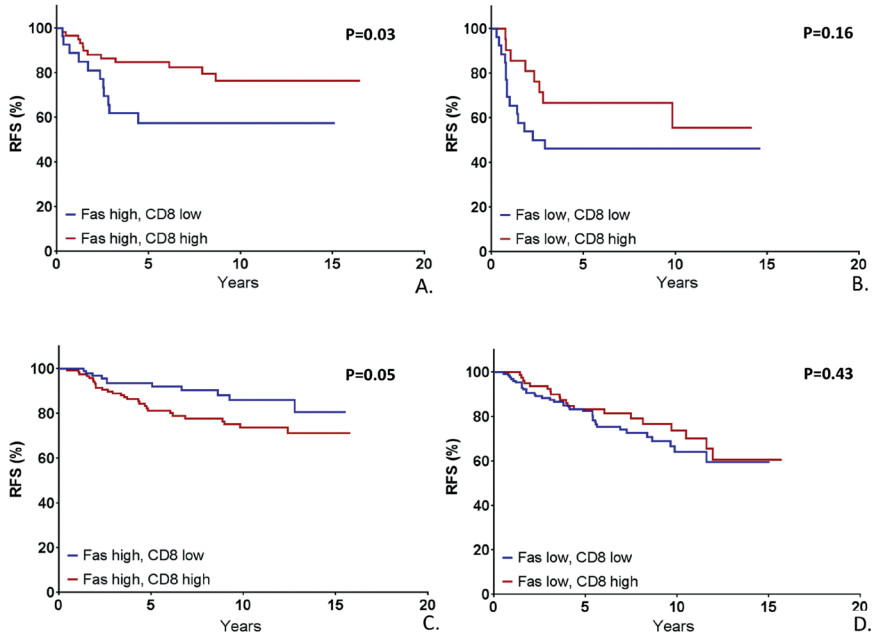
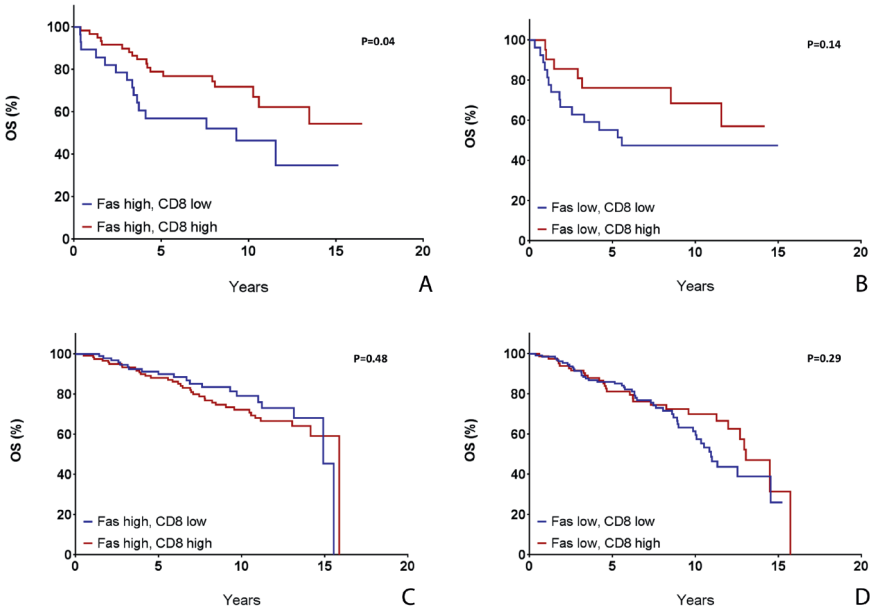


Figure 5 - Kaplan Meier survival estimates based on CD8+ tumor infiltrating lymphocytes (TILs) stratified on FAS expression in both ER-negative and ER-positive tumors. In patients with ER-negative tumors with high FAS expression, the high presence of CD8-positive is prognostic for a significantly higher survival (A). This prognostic effect is not observed in ER-negative tumors with low FAS expression (B), nor in ER-positive tumors (C, D). P-values represent log-rank survival test.



Supplementary Figure 3: The effect of CD8+ TILs on overall survival, stratified on both ER-status and FAS expression. In ER-negative tumors with high FAS expression, there is a significant benefit on OS of high CD8+ TIL presence, which is not observed in tumors with low FAS expression B. Furthermore, this effect is not shown in ER-positive tumors, either with high. C. or low D. FAS expression.

Discussion

This study is the first study to determine the influence of the immune-editing protein FAS on the prognostic value of TILs. Our studies shows that besides ER-negative status, also a positive FAS-status is required for CD8+ TILs to be prognostic in breast cancer.

Furthermore, we assess the value of FAS as an independent prognostic marker. It was shown that patients with a higher FAS-expression have a longer recurrence free and overall survival, even when corrected for age, histological subtype, tumor grade, tumor stage, ER-status, year of diagnosis, neo-adjuvant treatment, adjuvant chemotherapy and adjuvant endocrine therapy. This indicates that, irrespective of the amount of TILs, FAS serves as a prognostic marker. In our subgroup analysis, we confirmed that this effect was conserved in both ER-negative and ER-positive disease for recurrence free survival, but not for overall survival.

We also observed that FAS was expressed in nearly twice as much tumor cells in ER-negative tumors, compared to ER-positive tumors, which also explains the higher expression of FAS in younger patients, since those are more often ER-negative.[26] Even more, triple negative tumors had more than twice as much FAS-positive tumor cells compared to ER-positive tumors. This difference in FAS expression might explain the observation in earlier studies that infiltrating lymphocytes are only prognostic in ER-negative or triple negative tumors.[4] Since T-cell mediated immunity depends on FAS-expression of the target cell, a higher expression of FAS as observed in TNBC may render the cells more susceptible for infiltrating T-cells. With lower amounts of FAS-expression, as observed in ER-positive disease, infiltrating T-cells may have less possibilities to induce apoptosis, and will therefore be much less or even not prognostic since their functioning will be hampered. This explains why in ER-positive disease there is no additional value of CD8+ TILs over the prognostic value of FAS. However, we observed that in ER-positive tumors with high expression of FAS, there was even a significant negative effect of CD8-positive TILs on survival. This indicates that there is a FAS-independent, unknown factor which prevents CD8-positive TILs from functioning in ER-positive breast cancer Furthermore, since FAS itself is prognostic in ER-positive breast cancer, it indicates that there are more anti-tumor mechanisms of FAS expression besides the T-cell mediated immunity, which were described in detail earlier.[27] These mechanisms could have contributed to the favorable prognostic effect of FAS expression on clinical survival.

In ER-negative disease, it was observed that CD8+ TILs were only prognostic in the presence of high FAS expression, confirming the pivotal role of FAS in T-cell mediated immunity. Recently, immunotherapy has gained many interest in different fields of oncology, including breast cancer.[28, 29] These therapies are based on targeted agents (e.g. PD1 or PDL1 inhibitors), which enhance the immune response against the tumor.[30] We observed that TILs are only prognostic, and therefore perhaps are only functional, in the presence of FAS expression in ER-negative breast cancer. It could be hypothesized that FAS expression might therefore act as a predictive factor for these new emerging therapies.

Due to the retrospective, observational design of the cohort, there are some limitations to this study. First, HER2 was missing in nearly half of all patients based on the year of diagnosis, limiting the power for HER2-specific subgroup analyses). Therefore the year of diagnosis could have acted as a confounder, influencing both HER2-expression

and outcome. To overcome this confounding, the year of diagnosis was included as a corrective factor in multivariate analysis. Furthermore, since including HER2 in the multivariate model would lead to a skewed cohort only consisting of patients diagnosed after 2003, HER2 was not included as a corrective factor in multivariate analysis.

In summary, this study is the first study reporting a differential expression of FAS among tumors with different receptor subtypes; TNBC shows nearly twice as much expression compared to other subtypes. Furthermore, we showed that FAS is an independent prognostic marker in breast cancer, independent from estrogen receptor status or other possibly confounding factors. Finally, we showed that in ER-negative disease, FAS expression is necessary for the prognostic effect of TILs.

Materials & Methods

Patient cohort

The cohort of patients used for this study consisted of a consecutive series of female breast cancer patients treated in the Leiden University Medical Center (LUMC) with surgery between 1997 and 2009 (n=667). Data regarding age, year of diagnosis, estrogen and progesterone receptor expression, human epidermal growth factor receptor 2 (HER2) expression (when available), TNM stage[31], tumor differentiation grade[32] and morphology, local and systemic therapy, secondary tumors, local, regional, distant, recurrence free and overall survival time and status was known for these patients. Formalin-fixed paraffin-embedded (FFPE) tumor samples were collected, and a tissue microarray (TMA) was created with three 1mm tumor tissue punches from each tumor.

Immunohistochemistry

For immunohistochemical staining of FAS, 4.5 µm slides were cut from the aforementioned TMA and stored at +4 °C until use. Colon tissue was shown to be positive for FAS expression, therefore this tissue was used as positive control.[18] Slides were deparaffinized in xylene and rehydrated in serial dilutions of ethanol-H_2O. Antigen retrieval was performed by placing the section at 95°C for 10 minutes in Target Retrieval Buffer Low pH (DAKO) in a PT Link (DAKO). Endogenous peroxidase and phosphatase was blocked by incubation of the sections in BloxAll (Vector, Burlingame, USA) for 10 min. LS-B2820 (LifeSpan BioSciences, Seattle, USA) was used as anti-FAS

antibodies for IHC. The antibodies were diluted in phosphate-buffered saline with 1% of bovine serum albumin (1% PBS/BSA), and the optimal dilution was determined by titration. Incubations with primary antibodies were performed overnight. Envision HRP-labeled polymer anti-mouse (Dako, Carpinteria, USA) was used as secondary antibodies and incubated for 30 minutes. Slides were developed with DAB (Dako, Carpinteria, USA). Similar procedures were performed for a staining against CD8 to identify CD8+ cytotoxic T-cells (clone 4B11, Monosan).

In order to allow specific scoring of epithelial tumor cells, a counterstaining against stroma was performed using anti-rabbit polyclonal antibodies ab34710, ab6588 and ab23747 (Abcam, Cambridge, UK), targeting collagen I and IV and elastin respectively. Swine-anti-rabbit-AP (Dako, Carpinteria, USA) was used as secondary antibodies, incubated for 30 minutes and developed in the dark using VectorBlue Kit (Vector, Burlingame, USA). Finally, methyl green (Vector, Burlingame, USA) was used for staining of the nuclei. For this purpose, the section were incubated with methyl green for 5 min at 56 °C. After washing with demineralized water followed by acetone-HAc 0.05%, the sections were dehydrated by gradients of ethanol and dried by dipping in xylene. Slides were mounted in Vectormount (Vector, Burlingame, USA) and stored until further analysis.

Quantification of IHC stainings

The Philips Ultra Fast Scanner 1.6 RA (Philips, Eindhoven, the Netherlands) was used for digitalization of the immunohistochemically stained sections of the TMA. For FAS, the percentage of tumor cells showing membranous staining was assessed by two independent observers. The scores of the three punches were combined to one average score per patient. Based on the whole cohort, the median value was used as a cut-off value to create a dichotomous value distinguishing low and high expression of FAS. For the evaluation of CD8, the number of CD8+ cells in the tumor was counted per punch, and the average of three punches was used for dichotomization based on the median value. Punches were only analyzed when more than 30% consisted of tumor tissue. Images were acquired using a Leica ICC50 camera system (Leica Microsystems, Wetzlar, Germany).

mRNA expression analysis

To assess the correlation between FAS expression and ER-status at transcriptional level, the publicly available TCGA dataset was used using cBioPortal to assess the levels

of mRNA gene expression, in comparison to clinical ER-status and the expression of ESR1, the gene encoding for ER.[33, 34]

Statistical analysis

SPSS (version 23 for Windows) was used for statistical analysis. Chi-square, column proportion tests, Bonferroni's multiple comparisons test and unpaired t-tests were used to identify associations between FAS expression, CD8+ TIL presence and baseline clinicopathological parameters. Kaplan Meier analysis was used to calculate recurrence free (RFS) and overall survival (OS) for the complete cohort and subgroups; log-rank tests were used to assess any differences between survival curves. RFS was defined as the time without local, regional or distant recurrence, whereas OS was defined as death from any cause. Death from breast cancer (disease-specific survival) was not recorded for this cohort. Cox regression analysis was used for univariate and multivariate analyses for RFS and OS. Furthermore, interaction tests were performed, to assess the marker interaction effect. This test assesses whether the prognostic value of a marker in one subgroup is significantly different from its value in a different subgroup. For all tests, p-values <0,05 were considered to be significant.

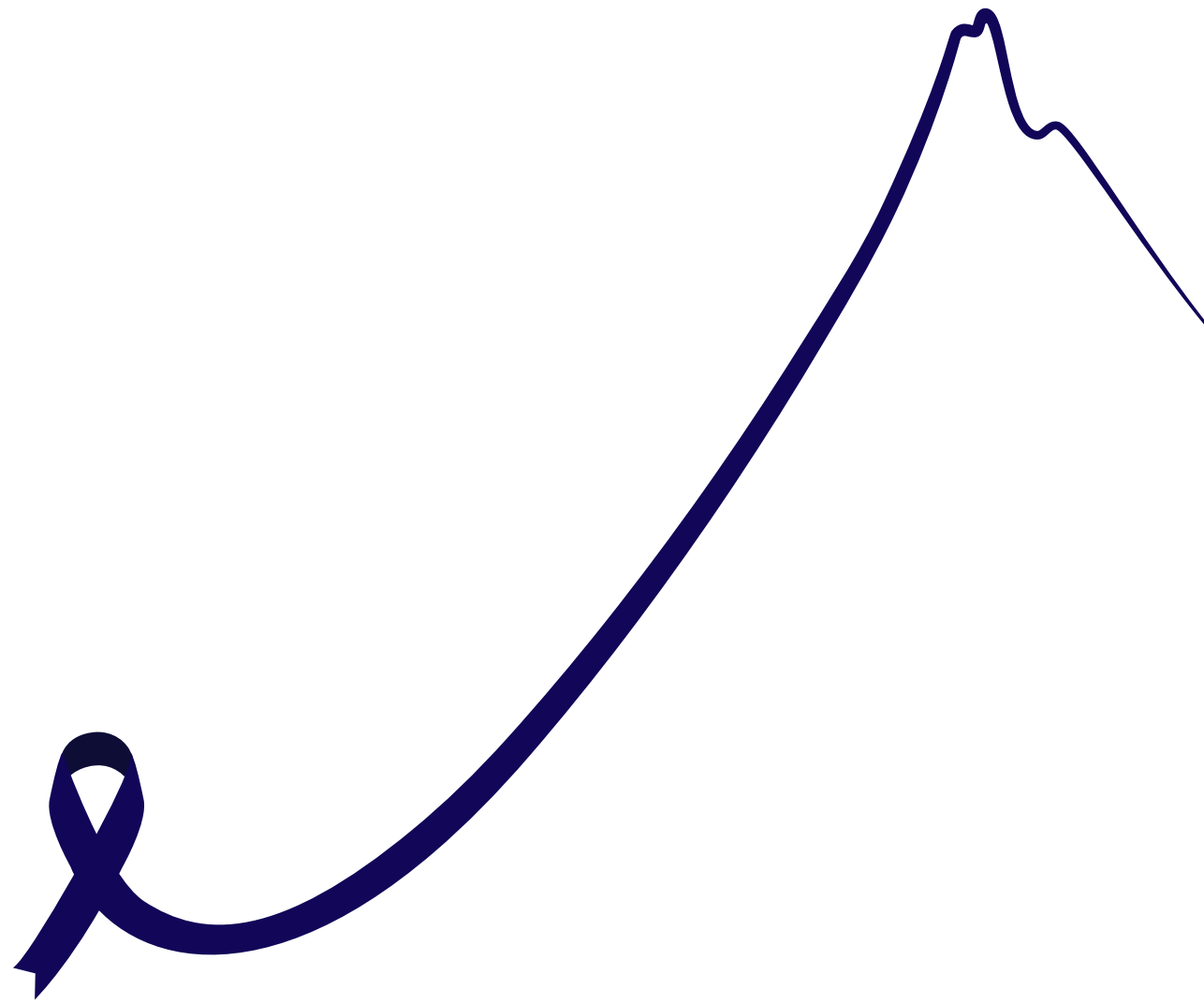
Acknowledgements

The authors sincerely thank all supportive staff in the research facilities of the department of Surgery of the LUMC. The results about FAS expression at transcriptional level are based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

References

1. Denkert C, Loibl S, Noske A et al. Tumor-Associated Lymphocytes As an Independent Predictor of Response to Neoadjuvant Chemotherapy in Breast Cancer. JCO 2010;28(1):105-113.
2. Whitford P, Mallon EA, George WD et al. Flow cytometric analysis of tumour infiltrating lymphocytes in breast cancer. Br J Cancer 1990;62(6):971-975.
3. Marrogi AJ, Munshi A, Merogi AJ et al. Study of tumor infiltrating lymphocytes and transforming growth factor-b as prognostic factors in breast carcinoma. Int J Cancer 1997;74(5):492-501.
4. Mahmoud SMA, Paish EC, Powe DG et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. JCO 2011;29(15):1949-1955.
5. West NR, Kost SE, Martin SD et al. Tumour-infiltrating FOXP3(+) lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. Br J Cancer 2013;108(1):155-162.
6. Liu S, Lachapelle J, Leung S et al. CD8(+)lymphocyte infiltration is an independent favorable prognostic indicator in basal-like breast cancer. Breast Cancer Res 2012;14(2):R48.
7. West NR, Milne K, Truong PT et al. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res 2011;13(6):R126.
8. Ali HR, Provenzano E, Dawson SJ et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. Annals of Oncology 2014;25(8):1536-1543.
9. Baker K, Lachapelle J, Zlobec I et al. Prognostic significance of CD8+ T lymphocytes in breast cancer depends upon both oestrogen receptor status and histological grade. Histopathology 2011;58(7):1107-1116.
10. Adams S, Gray RJ, Demaria S et al. Prognostic Value of Tumor-Infiltrating Lymphocytes in Triple-Negative Breast Cancers From Two Phase III Randomized Adjuvant Breast Cancer Trials: ECOG 2197 and ECOG 1199. JCO 2014.
11. Loi S, Michiels S, Salgado R et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Annals of Oncology 2014;25(8):1544-1550.
12. Loi S, Sirtaine N, Piette F et al. Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. JCO 2013;31(7):860-867.
13. Dieci MV, Mathieu MC, Guarneri V et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in two phase III randomized adjuvant breast cancer trials. Annals of Oncology 2015;26(8):1698-1704.
14. Ibrahim E, Al-Foheidi M, Al-Mansour M et al. The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. Breast Cancer Res Treat 2014;148(3):467-476.
15. Engels CC, Charehbili A, van de Velde CJH et al. The prognostic and predictive value of Tregs and tumor immune subtypes in postmenopausal, hormone receptor-positive breast cancer patients treated with adjuvant endocrine therapy: a Dutch TEAM study analysis. Breast Cancer Res Treat 2015;149(3):587-596.
16. Reimers M, Engels C, Putter H et al. Prognostic value of HLA class I, HLA-E, HLA-G and Tregs in rectal cancer: a retrospective cohort study. BMC Cancer 2014;14(1):486.
17. Krammer PH. CD95's deadly mission in the immune system. Nature 2000;407(6805):789-795.
18. The Human Protein Atlas. http://www.proteinatlas.org/ENSG00000026103-FAS/tissue. 2015. 18-12-2015. Ref Type: Internet Communication
19. Ashkenazi A, Dixit VM. Death Receptors: Signaling and Modulation. Science 1998;281(5381):1305-1308.
20. Waring P, Mullbacher A. Cell death induced by the Fas/Fas ligand pathway and its role in pathology. Immunol Cell Biol 1999;77(4):312-317.
21. Reimer T, Herrnring C, Koczan D et al. FasL:Fas Ratio: Prognostic Factor in Breast Carcinomas. Cancer Research 2000;60(4):822-828.
22. Sjostrom J, Blomqvist C, von BK et al. The predictive value of bcl-2, bax, bcl-xL, bag-1, fas, and fasL for chemotherapy response in advanced breast cancer. Clin Cancer Res 2002;8(3):811-816.
23. Mottolese M, Buglioni S, Bracalenti C et al. Prognostic relevance of altered Fas (CD95)-system in human breast cancer. Int J Cancer 2000;89(2):127-132.

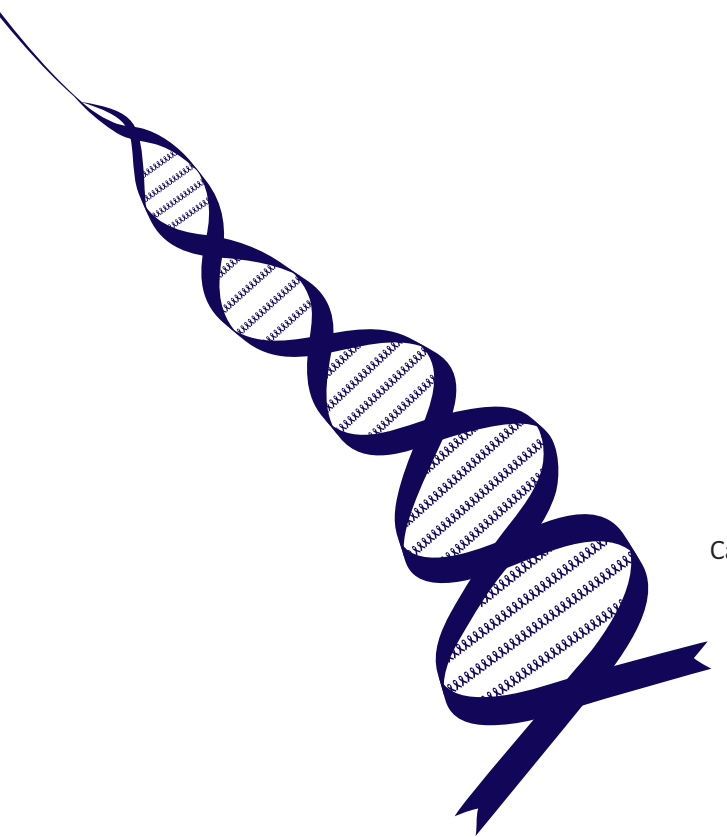
24. Debatin KM, Krammer PH. Death receptors in chemotherapy and cancer. Oncogene 0 AD;23(16):2950-2966.
25. Comprehensive molecular portraits of human breast tumours. Nature 2012;490(7418):61-70.
26. Lee AJ, Cunningham AP, Kuchenbaecker KB et al. BOADICEA breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface. Br J Cancer 2014;110(2):535-545.
27. Peter ME, Hadji A, Murmann AE et al. The role of CD95 and CD95 ligand in cancer. Cell Death Differ 2015;22(4):549-559.
28. Nanda R, Chow LQM, Dees EC et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. JCO 2016.
29. Mittendorf EA, Philips AV, Meric-Bernstam F et al. PD-L1 Expression in Triple-Negative Breast Cancer. Cancer Immunology Research 2014;2(4):361-370.
30. Topalian SL, Drake CG, Pardoll DM. Targeting the PD-1/B7-H1 (PD-L1) pathway to activate anti-tumor immunity. Current opinion in immunology 2012;24(2):207-212.
31. Frederick L, Page DL, Fleming ID et al. AJCC cancer staging manual. 1 ed. Springer Science & Business Media; 2002.
32. Bloom HJG, Richardson WW. Histological Grading and Prognosis in Breast Cancer: A Study of 1409 Cases of which 359 have been Followed for 15 Years. Br J Cancer 1957;11(3):359-377.
33. Cerami E, Gao J, Dogrusoz U et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery 2012;2(5):401-404.
34. Gao J, Aksoy BA, Dogrusoz U et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci Signal 2013;6(269):l1.



Chapter 9

Systematic review of the clinical and economic value of gene expression profiles for invasive early breast cancer available in Europe

E.J. Blok
E. Bastiaannet
W.B. van den Hout
G.J. Liefers
V.T.H.B.M. Smit
J.R. Kroep
C.J.H. van de Velde



Cancer Treat Rev, 2018 Jan; 62: 74-90

Abstract

Gene expression profiles with prognostic capacities have shown good performance in multiple clinical trials. However, with multiple assays available and numerous types of validation studies performed, the added value for daily clinical practice is still unclear. In Europe, the MammaPrint, OncotypeDX, PAM50/Prosigna and Endopredict assays are commercially available. In this systematic review, we aim to assess these assays on four important criteria: Assay development and methodology, clinical validation, clinical utility and economic value.

We performed a literature search covering PubMed, Embase, Web of Science and Cochrane, for studies related to one or more of the four selected assays.

We identified 147 papers for inclusion in this review. MammaPrint and OncotypeDX both have evidence available, including level IA clinical trial results for both assays. Both assays provide prognostic information. Predictive value has only been shown for OncotypeDX. In the clinical utility studies, a higher reduction in chemotherapy was achieved by OncotypeDX, although the number of available studies differ considerably between tests. On average, economic evaluations estimate that genomic testing results in a moderate increase in total costs, but that these costs are acceptable in relation to the expected improved patient outcome. PAM50/prosigna and EndoPredict showed comparable prognostic capacities, but with less economical and clinical utility studies. Furthermore, for these assays no level IA trial data are available yet.

In summary, all assays have shown excellent prognostic capacities. The differences in the quantity and quality of evidence are discussed. Future studies shall focus on the selection of appropriate subgroups for testing and long-term outcome of validation trials, in order to determine the place of these assays in daily clinical practice.

Keywords: Breast cancer, gene expression, OncotypeDX, MammaPrint, Prosigna, Endopredict

Conflicts of interest: none

Introduction

In the past decades, there has been a steady increase in the survival rates of patients with breast cancer. Among other factors like early screening and awareness, the majority of this effect is attributed to the concept of adjuvant therapy.[1,2] However, among all patients receiving adjuvant chemotherapy, the majority would not have developed metastases even without adjuvant therapy, whereas in contrast some patients without the indication for adjuvant therapy still develop distant metastases. A recent progress in this optimal selection is the development of genomic profiling assays.[3] We chose four crucial criteria for determining the value of these assays.

Assay development and methodology

The first criterion is the methodological robustness, both during development and during the commercial activities. For example, the tests should be validated in a cohort independent from the training cohort, and should not be used in a patient population in which the test was not validated unless re-validation is performed. Furthermore, there should be little to no inter-test variation when the same tissue samples are tested multiple times.

Another aspect of assay development is determining the target population. Therefore, studies need to focus on identifying subgroups which do not benefit from genomic testing since the outcome of the test overlaps with the stratification by the clinicopathological factors (e.g. when all or almost all triple-negative breast cancers are considered high-risk by the test).

Clinical validation

A second important factor is the effect on clinical outcome between the different test-outcome groups. Similar to classical biomarkers, a distinction can be made between the prognostic and the predictive value of a test.[4] Since the utility of genomic testing is in particular aimed at guiding decisions regarding chemotherapy, a predictive test, able to predict which patients will benefit from chemotherapy or not, is more valuable than a solely prognostic test which is only associated with the patient prognosis.

Clinical utility

The third criterion is the clinical utility of the test. Applying the test should lead to a shift in the indication of chemotherapy as compared to indication based on traditional

parameters. In other words, if the patients using chemotherapy based on the test results are exactly the same patients as the ones using chemotherapy based on the traditional clinicopathological parameters, the test has no additional value.

Economic value

The fourth, and last criterion for genomic testing is the economic value of the test. Due to the commercialisation of the assays, the tests are more expensive than the regular pathological assessment, with costs ranging from €1800 to €3700 per test. In an era of emphasis on healthcare efficiency, the costs of the test should be justified by its clinical and health benefits, and the reduction in costs by reducing adjuvant therapy use.

Test descriptions

The first test, which was first developed in 2002 by van 't Veer et al and for which the prognostic capacities were shown simultaneously by van de Vijver et al, is the 70-gene prognosis profile, better known as MammaPrint (Agendia BV, Amsterdam, The Netherlands).[5,6] This assay uses the mRNA expression of 70 genes using microarray technology, to categorize patients in either a low or high risk . These 70 genes were identified from a total of 25,000 genes using supervised clustering.

The second test in this review is the 21-gene Recurrence Score, also known as the OncotypeDX Recurrence Score (RS) (Genomic Health Inc., Redwood City, CA). The test is based on the expression of 21 genes in FFPE cancer tissue, determined using reverse transcriptase PCR (RT-PCR)[7]. Of these genes, 16 genes are cancer-related and were selected out of 250 rationally selected candidate genes based on their prognostic capacity and consistency in test performances.[7] Based on these relative expressions, the Recurrence Score is calculated ranging from 0 to 100, with low risk ranging from 0 to 17, intermediate risk ranging from 18 to 30, and high risk ranging from 31 to 100. However, for the most important validation trial of this test, the risk categories in this trial were adjusted to 0-10, 11-25 and 26-100 for the low-, intermediate- and high risk respectively.[8]

The third test included in this review is the Prosigna, based on the better-known PAM50 test (NanoString Technologies, Seattle, WA). This test, based on the expression of 46 genes using quantitative PCR (qPCR) is able to distinguish between the molecular subtypes of breast cancer (luminal A, luminal B, HER2-enriched, normal-

like and basal-like).[9] Furthermore, it provides the risk of recurrence score (ROR) and the subsequent risk category. The test was adapted by NanoString in order to allow the use in local pathology laboratories.[10]

The fourth and last test which will be discussed in this systematic review is the EndoPredict (Myriad Genetics Inc, Salt Lake City, UT). This assay uses the expression of 8 cancer-related and 3 reference genes determined by RT-PCR, which results in a risk score from 0 to 15 (EP), which is subsequently divided into low and high risk.[11] A special feature of the EndoPredict is the integration of tumour size and nodal status, resulting in an EP clinical score (EPclin). The EndoPredict can be performed in local laboratories, in contrast to the MammaPrint and OncotypeDX which are centrally determined and therefore need more elaborate logistical planning.

In this review, we evaluate four genomic assays available in Europe using a systematic evaluation focusing on all four major criteria with the aim to assess each test individually for its strengths and weaknesses.

Methods

Search strategy

This systematic review was to comprehensively cover all four aspects of the four commercially available genomic profiling tests in Europe on four different aspects: developmental and methodological robustness, extend of clinical validation, clinical utility and economic value. These items were chosen after a consensus meeting and cover those evaluation criteria we deemed most important. We searched PubMed, Embase, Web of Science and Cochrane for articles published before April 2016. The search strategy (supplementary document 1) was applied on April 7th 2016, and after evaluation of all abstracts it was updated at September 9th 2016. Abstracts were screened for relevance based on the title and abstract, and remaining full-text articles were screened based on the inclusion criteria.

Selection criteria

Articles were selected if they studied one of the four tests available in Europe: OncotypeDX, MammaPrint, Prosigna or Endopredict. Furthermore, the article should be original peer-reviewed research; abstracts, posters, reviews and meta-analyses

were excluded. The article needed to cover one of the four criteria: development of the test, clinical validation, clinical utility or an economical evaluation. For the clinical validation studies, survival analysis was required, evaluating either the differences in survival between test-outcome groups, or the benefit of therapy in one or more test-outcome groups. For the clinical utility studies, decision impact studies were to be available in a representative cohort, and had to report both the absolute increase or decrease in chemotherapy as well as the shift from one treatment category to the other. Retrospective large-scale population-based impact studies were also included, reporting real-life shifts in the use of genomic testing and the subsequent changes in therapy decisions. Two reviewers (EJB, EB) independently selected articles that met the above inclusion criteria based on title and abstracts. Next, full-texts of potentially relevant articles were screened. Agreement concerning eligibility was achieved during consensus.

Data extraction and statistics

Data extraction was independently performed by the two reviewers. Data was collected concerning the performed test, the number of included patients, the results of the test, and survival outcome or change in treatment where appropriate. Disagreements in data extraction and interpretation were resolved during a consensus meeting. There were no changes in eligibility criteria during the selection of articles. All studies that fulfilled the inclusion criteria were included, independent of their methodological quality; no risk of bias assessment was performed. Both retrospective and prospective studies were included without exclusion of particular study designs with an emphasis on prospective RCTs (where available). Data were recorded in the tables as mentioned in the articles, no additional statistics were performed. Both point estimates and 95%CI were recorded, where appropriate and mentioned in the selected articles.

Due to the heterogeneity of the studies chosen, the patient selection and endpoints reported, no further statistical analyses could be performed. Results were stratified in (1) one of the four tests and (2) lymph node positive or lymph node negative patients or articles where the distinction could not be made or both groups were included.

For the clinical utility, extracted data from decision-impact studies were pooled (weighted by the number of patients) to give an estimate of the chemo-reduction and shift in therapy a test can establish. We only considered a change in chemotherapy and recorded the percentage of patients who would receive chemotherapy before

the test, and after the test (as mentioned in the included articles). For the table on clinical validation, the number of patients who were high or low risk according to the test were recorded and the outcome in the groups. Outcomes were recorded as mentioned in the articles: distant metastasis or distant recurrence free survival, breast cancer specific survival, and overall survival were most frequently reported. Where known, both the point estimate and the 95%CI were recorded. The Hazard Ratio and corresponding 95%CI for the difference in outcome between the risk groups was recorded if this was mentioned in the articles. For the economic review, original evaluations were included if they compared costs beyond the assay costs alone. Evaluations could be cost minimization analyses (CMA), cost effectiveness analyses (CEA, comparing costs to life years) or cost utility analyses (CUA, comparing costs to quality-adjusted life years (QALYs)). To aggregate, QALYs were imputed for CMAs and CEAs (as predicted by the average and the life year gain, respectively) and costs were updated to Euros at price level 2016. When more than one (non) genomic strategy was included in an economic evaluation, the (non-) genomic strategy with the highest QALYs was used in the review.

Results

Using our search strategy, we identified 1345 unique titles and abstracts. Limiting ourselves to the manuscripts only related to the topics of this review, we selected 280 studies for further full-text evaluation. From these 280 full-text manuscripts, we selected 149 papers for inclusion in this review: 11 about developmental validation, 12 about biomarker prediction, 50 about clinical studies, 28 about clinical utility and the effect on chemotherapy reduction, 44 economic evaluations and 4 studies making direct head-to-head comparisons on test outcome between two or more of the included tests (figure 1).

Assay development and methodology

In the development of MammaPrint, multiple evolutions were necessary to allow high-throughput screening of FFPE tissue. Glas et al first converted the original research-based micro-array containing approximately 25,000 probes to a mini-assay with good concordance and reproducibility.[12,13]

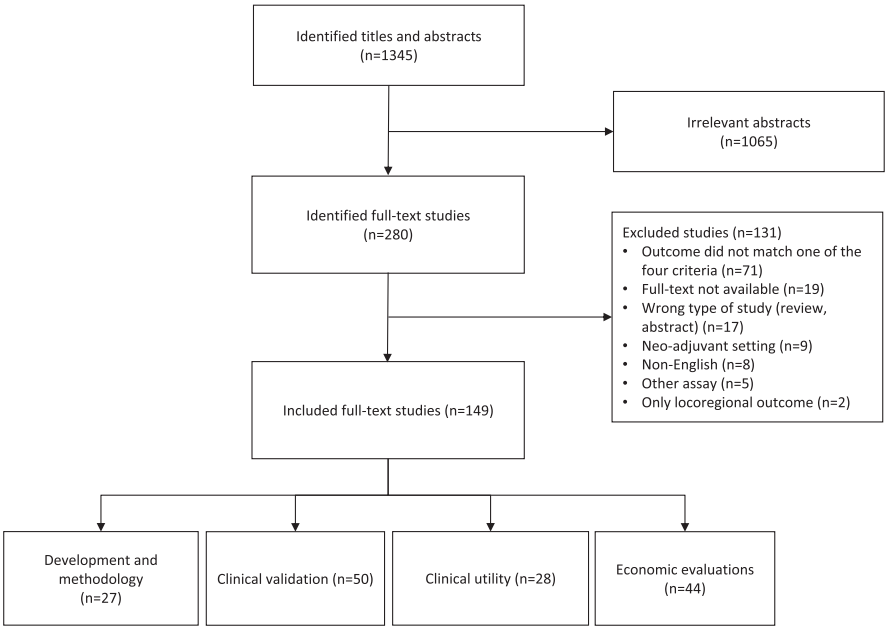


Figure 1. A diagram showing the inclusion of relevant papers in the systematic review

A second step was the conversion from frozen to FFPE tissue by Mittempergher et al, with an R^2 of 0.94.[14] After this proof of principle, Sapino et al further developed the MammaPrint towards an FFPE platform, again with a good correlation between FFPE and frozen tissue (r=0.92), and a high concordance between high- and low-risk classifications between both methods (κ-score 0.82).[15] Beuner et al validated both the conversion to a mini-assay and the conversion from frozen tissue to FFPE retrospectively, by comparing the scores of both methods.[16]

Gyanchandani et al studied whether intratumoral heterogeneity might influence the outcome of a gene expression test in 74 ER-positive cases using most included gene expression panels, by assessing different tumor regions from the same FFPE block.[17] They showed that genomic assays with a higher number of included genes resulted in a lower rate of discordant samples. Drury et al studied the use of 0.6mm cores and compared these with full sections, to establish whether tissue-microarrays (TMAs) could be used for genomic profiling using OncotypeDX.[18] Although the total RNA yield was lower from tissue cores compared to full sections, the OncotypeDX Recurrence Score results from individual cores clustered closely, and had an excellent correlation with full-section RS (Spearman R=0.91).

For the Endopredict, the use of pre-surgery biopsies and surgical sections from 40 ER-positive HER2-negative tumors was compared. It was shown that comparing both results resulted in a Pearson correlation coefficient of 0.92, showing that core needle biopsies can be used for genomic profiling using Endopredict.[19] Another aspect of the EndoPredict is decentral assessment, meaning that every individual pathological laboratory can perform this test and thereby reducing the logistical strain on the testing procedure. Denkert et al tested this decentral evaluation.[20] The Pearson correlation coefficient for all measurements was a near-perfect 0.994, and 100% of the samples were assigned to the same EP risk group as the reference test. Furthermore, Kronenwett et al showed that this decentral approach had excellent precision and reproducibility, although with a small sample size.[21]

Although these published studies showed a good reliability and reproducibility, the MINDACT trial shows that there can be problems which hamper the reliability and feasibility of a test. Between May 2009 and January 2010, 162 patients were falsely identified as being high risk, due to a change in RNA-extraction solution.[22] Furthermore, of all 11,288 screened patients, there was a screening failure in 1182 patients (10%) in which the MammaPrint was not feasible.[22]

Another concern for the reliability of test results is the ratio between tumor and normal tissue in the tested specimen. Elloumi et al showed that an increase of normal tissue in the specimen leads to biased test results when compared to uncontaminated tumor tissue test results.[23] For the PAM50 this bias was linear, showing a more favourable outcome with increasing normal tissue content. For the MammaPrint and OncotypeDX the bias was unpredictable, switching both from low to high risk and vice versa with increasing normal tissue content. All tests have since developed strategies to mitigate this bias.

A couple of studies directly compared the test results of multiple tests performed on one tumour. In the OPTIMA Prelim trial, patients were randomized between standard therapy or OncotypeDX-directed therapy.[24,25] Among others, also MammaPrint and Prosigna tests were performed. Strikingly, the kappa measurements were between 0.40 and 0.53. In the same cohort of patients, OncotypeDX predicted 17.9% to be high risk, compared to 38.6% and 34.5% for MammaPrint and Prosigna respectively. This pilot trial is now followed by the OPTIMA trial, in which treatment directed by the Prosigna assay is compared with regular care. In a smaller prospective study, 52

samples were analysed with both the OncotypeDX and Prosigna, showing a Spearman correlation coefficient of 0.08.[26] Remarkably, 57.1% of the patients classified as high risk by Prosigna were classified as low risk by OncotypeDX. In a similar study comparing Endopredict and OncotypeDX results in 34 samples, a Pearson correlation of 0.65 was shown, with a concordance between risk categories of 76%.[27]

Prediction of test results

Theoretically, a genomic profile can have an excellent prognostic value, but is 100% predicted by the occurrence of other markers and therefore has no added value. Therefore, it is crucial to establish the added value of the test, by testing whether the test result can be predicted by standard clinicopathological parameters. This testing could identify subgroups for which the test is not valuable. We identified 12 studies evaluating this effect, which are reported in table 1. In general, tumours which are (a combination of) grade 1, PR-positive and/or have a Ki-67 expression lower than 10%, are almost always low risk when genomic testing is performed. Similarly, tumours which are (a combination of) grade 3, PR-negative and/or have a Ki-67 score of more than 40%, are almost always high-risk. For these subgroups, genomic profiling provides little additional information.

Table 1 Marker prediction, according to test and nodal status

Marker prediction

Authors	Year	Patients (N)	Markers in best-fit model	R^2 best fit model	Subgroups little/no benefit of testing (>75% in risk category)	
MammaPrint						
Early stage breast cancer (combined LN- and LN+, other groups or not specified)						
Cardoso*[22]	2016		NA	NA	Grade 1 Grade 3 ER- PR-	93% low risk 75% high risk 96% high risk
Gevensleben[28]	2010	140	NA	NA	St. Gallen high risk St Gallen low risk Grade 1 Grade 3 PR-negative	80% high risk 86% low risk 79% low risk 76% high risk 76% high risk"
OncotypeDX						
Lymph node negative						
Chaudhary[29]	2016	350	NA	NA	PR+	95% low or intermediate risk
Dialani[30]	2016	319	ER, PR, HER2, tumor grade	0.55	NA	NA
Sparano*[8]	2015	8523	NA	NA	PR- Grade 3	5% low risk 11% low risk
Ingoldsby[31]	2013	52	PR (allred), nuclear pleomorphism (np), survivin	NA	Grade 1 PR <2, np-score 3	100% low or intermediate risk 100% high risk
Sahebjam[32]	2011	53	PR, Ki-67	0.84	Ki-67 <10%	100%= low or intermediate risk
Auerbach[33]	2010	138	Mitotic count, PR	NA	PR+ &Mitotic count 1 or 2 PR- & Mitotic count 2 or 3	100%= low or intermediate risk 75%= high risk (0% low risk)
Flanagan[34]	2008	42	ER, PR, grade, HER2, mitotic count	0.66	Grade 1 Grade 3	100% low or intermediate risk 83.3% high risk (0% low risk)
Wolf[35]	2008	300	NA	NA	PR+ & Grade 1/2	94% low or intermediate risk
Early stage breast cancer (combined LN- and LN+, other groups or not specified)						
Gluz[36]	2016	2642	NA	NA	Grade 1 Ki-67 <20%, PR >20% Ki-67 >40%	~90% low or intermediate risk ~95% low or intermediate risk ~90% high risk
Bradshaw[37]	2013	158	ER (allred), PR (allred), Ki-67	0.62	NA	NA
Allison[38]	2012	173	PR, tumor grade	Unknown (p<0.001)	Grade 1 & PR >5 (allred) Grade 3 & PR <5 (allred)	100% low or intermediate risk 80% high risk (0% low risk)
Williams[39]	2011	133	NA	NA	Ki-67 <10%	99%= low or intermediate risk

*Not designed to predict test results, but data are provided in the manuscript

Table 2 Clinical validation, according to test and patient inclusion

Clinical validation

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
MammaPrint							
Lymph node negative							
Wittner[40]	2008	D	N=100, LN-, postmenopausal	Low 27 High 73	NPV DM 100% (87-100) PPV DM 12% (6-22)		
Van 't Veer[5]	2002	D	N=78, LN-, <5cm, <55 years	Low 44 High 34	PPV DM 83% Validation set 2/19 incorrect		
Mook[41]	2010	D	N=148, T1-2N0, 55-70 years	Good prognosis 91 Poor prognosis 57	DMFS 93% (SE 3%) vs 72% (SE 6%) BCSS 99% (SE 1%) vs 80% (SE 5%)	DMFS 4.6 (1.8-12.0; p=0.001) BCSS 2.0 (1.0-4.0; p=0.04)	Adjuvant! 62 (42%) good prognosis; 45 (30%) poor prognosis
Na[42]	2011	D	N=36, cT1-2N0M0	Low 5 High 31	40% low-risk prognosis, 60% high-risk prognosis		St. Gallen guidelines 29 (80.6%); 30 (83.4%) NIH guidelines; 23 (63.8%) Adjuvant! Online
Drukker[43]	2013	C	N=427, cT1-3N0M0	Low 219 High 208	5-yrs DRFS 97.0 5-yrs DRFS 91.7		
Bueno de-Mesquita[44]	2007	C	N=427, cT1-4N0M0, <61 years	Good 219 Poor 208			Discordance Adjuvant! 160 (37), St Gallen 168 (39), NPI 117 (27)
Bueno de-Mesquita[45]	2009	D	N=123, pT1-2N0, <55 years	Good 52% Poor 48%	DMFS 98(±2) good vs 78(±6) poor OS 97(±2) good vs 82(±5) poor	5.7 (1.6-20) 3.4 (1.2-9.6)	
			Second series N=151	Good 40% Poor 60%	DMFS 86(±5) vs 50(±6) OS 94(±3) vs 51(±5)	5.5 (2.5-12) 10.7 (3.9-30)	
Lymph node positive							
Knauer[46]	2010	D	N=541 LN+, received ET or ET+CT	Low 252 High 289	BCSS 97% low vs 87% high DMFS 95% low vs 82% BCSS low: 97% ET, 99% ET+CT BCSS high: 81% ET, 94% ET+CT DDFS low: 93% ET, 99% ET+CT DDFS high: 76% ET, 88% ET+CT	4.81 (1.98-11.67) 3.88 (1.99-7.58) 0.58 (0.07-4.98) 0.21 (0.07-0.59) 0.26 (0.03-2.02) 0.35 (0.17-0.71)	

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Saghatchian[47]	2012	D	N=173, 4-9 positive lymph nodes	Low 70 High 103	OS 97% vs. 76% high risk DMFS 87% low vs 63% high (p < 0.01)	HR 2.211, p=0.005	
Mook[48]	2009	D	N=241, T1-3, 1-3 positive LN	Good 99 Poor 142	DMFS 91% (SE 4%) vs 76% (SE 4%) BCSS 96% (SE 2%) vs 76% (SE 4%)	4.13 (1.72-9.96; p=0.002) 5.70 (2.01-16.23; p=0.001)	Discordant Adjuvant! 77 (32%)
Early stage breast cancer (combined LN- and LN+, other groups or not specified)							
Mook[49]	2010	D	N=964, pT1	Good 525 Poor 439	DMFS 87% SE 2% vs 72% SE 3%, BCSS 91% SE 2% vs 72% SE 3%	DMFS 2.70 (1.88-3.88) BCSS 4.22 (2.70-6.60)	
Van de Vijver[6]	2002	D	N=295, stage I-II, <53 years	Poor 180 Good 115	Mean OS 54.6(±4.4) vs 94.5(±2.6) DMFS 85.2(±4.3) vs 50.6(±4.5)	5.1 (2.9-9.0)	
Knauer[50]	2010	D	N=168, T1-3N0-1, HER2+	Good 20 Poor 69	DMFS 84% vs DMFS 55% No data BCSS	DMFS 4.5 (1.1-18.7) BCSS 3.8 (0.9-15.8)	
Drukker[51]	2014	D	N=295, T1-2N0-1M0, <53 years	Low 115 High 180	25-yrs DMFS 60.4 (45.3-80.5) vs 41.6 (32.6-53.1) 25-yrs OS 57.3 (44.8-73.2)vs 39.7 (31.7-49.8)	DMFS 3.1 (2.02-4.86) OS 2.9 (1.90-4.28)	
Drukker[52]	2014	D	N=1053, T1-3N0-1M0	Low 561 High 492	LRR 6.1 (4.1-8.5) LRR 12.6 (9.7-15.8)	2.40 (1.54-3.74)	
Cardoso[22]	2016	A	N=6693 early stage BC	Low CR-low GR 2745, Low CR-High GR 592, High CR-Low GR 1550, High CR-High GR 1806	Chemo: 95.8 (92.9-97.6) No chemo: 95.0 (91.8-97.0) DMFS No chemo: 94.4 (92.3-95.9) DMFS chemo: 95.9 (94.0-97.2)	CT vs no CT: 1.17 (0.59-2.28) CT vs no CT: 0.78 (0.50-1.21)	14.3%

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Buyse[53]	2006	D	N=307	Clinical low risk GLR 52 / GHR 28 Clinical high risk GLR 59 / GHR 163	10-yrs OS 0.88 / 0.69 0.89 / 0.69	DMFS 2.32 (1.35-4.00) vs Adjuvant! 1.68 (0.92-3.07) OS: 2.79 (1.60-4.87) vs 1.67 (0.93-2.98)	
Kunz[54]	2011	D	N=689, 35-55 years	Low 42% High 58%	10-yrs OS: 90.2% (86.3-94.1) 65.2% (60.3-70.1) 10-yrs DMFS 87.7 (83.6-91.8) vs 64.5 (59.8-69.2)		St. gallen low risk 4, high risk 6, intermediate 34
Kok[55]	2012	D	121 with TAM, 151 no TAM, 92 TAM for M1	Good 83 / Poor 38 Good 85 / Poor 66 Good 45 /Poor 47	BCSS 80.6 (SE 4.7) vs 63.4 (SE 8.2) BCSS 90.2 (3.3) vs 63.3 (6.3) Median TTP 20.9 vs 6.6 months	2.78 (1.30-5.94) 4.52 (2.01-10.2) 2.55 (1.59-4.07)	
Ahn[56]	2013	C	N=82, ER+, with intermediate score Oncotype	Good 66 Poor 16		Multivariable 10.19 (1.05–99.01); P=0.013	
Ishitobi[57]	2010	D	N=102, <70 yrs,	Low 20 High 82	DMFS 100% DMFS 94%		PPV 9.8%, NPV 100%
OncotypeDX							
Lymph node negative							
Toi[58]	2010	D	N=200, N0, ER+	Low 48% Intermediate 20% High 33%	DRFS 3.3 (1.1-10.0) vs 0% vs 24.8 (15.7-37.8) OS 6.4 (2.9-13.6) vs 2.6 (0.4-16.8) vs 19.1 (11.3-31.3)	HR for 50-point increase 6.09 (2.17-16.7), p<0.001	
Paik[7]	2004	D	N=668, N0, ER+	Low 51% Intermediate 22% High 27%	DRFS 6.8 (4.0-9.6) vs 14.3 (8.3-20.3) vs 30.5 (23.6-37.4)	3.21 (2.23-4.61); p<0.001	
Naoi[59]	2013	D	N=459, N0, ER+	Low 286 Intermediate 81 High 92	RFS better low vs intermediate p=0.0014 and high P = 1.7e-11		

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Sparano[8]	2015	A	N=10253, N0, HR+, HER2-, 1.1-5.0 cm	Low 1629 Intermediate 6907 High 1736	RS 0-10: 5-yrs invasive DFS 93.8 (92.4-94.9); DRFS 99.3 (98.7-99.6); RFS 98.7 (97.9-99.2); OS 98.0 (97.1-98.6)		
Mamounas[60]	2010	B	N=895, N0, ER+	Low 862 Intermediate 368 High 444	10-yr LRR 4.3 (2.3-6.3) low, 7.2 (3.4-11.0) intermediate, 15.8 (10.4-21.2) high RS Placebo: p=0.022: 10.8% (5.8% to 15.8%) vs 20.0% (9.9% to 30.0%) vs 18.4% (9.5% to 27.4%) CT + TAM: p=0.028; 1.6% (0% to 3.5%) vs 2.7% (0% to 6.4%) vs 7.8% (2.6% to 13%)	HR 2.16 for 50 units in RS (1.26 to 3.68; P=0.007)	
Sgroi[61]	2013	B	N=665, N0, ER+	BCI vs OncotypeDX vs IHC4	-Early DR: BCI HR 2.77 [95% CI 1.63–4.70], LR-$\Delta\chi^2$=15.42, p<0·0001; 21-gene recurrence score HR 1.80 [1.42–2.29], LR-$\Delta\chi^2$=18.48, p<0·0001; IHC4 HR 2.90 [2.01–4.18], LR-$\Delta\chi^2$=29.14, p<0.0001 -Late DR BCI HR 1.95 [95% CI 1.22–3.14], LR-$\Delta\chi^2$=7.97, p=0.0048; 21-gene recurrence score HR 1.13 [0.82–1.56], LR-$\Delta\chi^2$=0.48, p=0.47; IHC4 HR 1.30 [0.88–1.94], LR-$\Delta\chi^2$=1.59, p=0.20)		
Paik[62]	2006	B	N=651, N0, ER+	Low 353 Intermediate 134 High 164	DR: Low chemotherapy RR 1.31 (0.46-3.78), intermediate RR 0.61 (0.24-1.59); High RR 0.26 (0.13-0.53)	Test interaction chemotherapy and RS p=0.038	
Tang[63]	2011	B	N=1444 N0, ER+	Intermediate RSPC (17.8%)	DR vs RS RSPC vs RS: intermediate 17.8 vs 26.7, and low risk 63.8 vs 54.2	Interaction term RSPC chemotherapy p=0.10	

Table 2 continued

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Yorozuya[64]	2010	D	N=40, N0, ER+, Stage I-IIA	Cases 10, controls 30. Cases: low 3, intermediate 1, high 6; Controls low 19, interm 8, high 3	Mean RS cases 40.0 (21.1-58.9), controls 17.8 (13.8-21.9); p<0.001		
Lymph node positive							
Albain[65]	2010	B	N=367, N+, ER+, postmenopausal	Low 146 Intermediate 103 High 118	DFS TAM alone HR 2.64 (1.33-5.27) 50point difference. Benefit chemotherapy low risk: HR 0.97 (0.54-1.93), high risk HR 0.59 (0.35-1.01)	Interaction RS treatment p=0.029 1st 5 yrs, beyond 5yrs p=0.58	
Early stage breast cancer (combined LN- and LN+, other groups or not specified)							
Gluz[36]	2016	A	N=348, pN0-1, HR+	Low risk 18.1% Intermediate 60.4 High 21.6%	3-yrs DFS 98% RS<=11, 96% RS 12-25, 92% RS>25		
Dowsett[66]	2010	B	N= 1231, HR+, postmenopausal	N0: 59, 26, 15% N+: 52, 31, 17	9-yr DR 4%, 12% and 25% in N0 9-yr DR 17%, 28% and 49% in N+	RS 50 units: N0: 5.25(2.84-9.73), N+ 3.47 (1.64-7.38)	
Cuzick[67]	2011	B	N= 1125, ER+		DR RS vs IHC4 score (r=0.68 for TTDR, N0; r=0.71 for time to recurrence [TTR], all; r=0.69 TTR,N0)		r=0.72
Goldstein[68]	2008	D	N=465, HR+, 0-3 positive nodes	Low 46% Intermediate 30% High: 24%	5-yr recurrence rate low risk: 0-1 positive nodes 3.3 (2.2-5.0), 2-3positive nodes: 7.9% (4.3-14.1)	Multivariable – 50-point increase) 2.12 (0.97-1.65)	
Le Du[69]	2015	D	N=1030, stage I, ER+, HER2-	Low 571 Intermediate 370 High 89	DDFS: RS predictor DDFS p<0.001. High RS 76.4% (59.2-87.1), low RS 95.9%(93.0-97.6)	HR 2.197 (0.901-5.356), p=0.083	

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Tang[70]	2010	B	N=668	Low: 338 Intermediate 149 High 181	Adjuvant! Low: DRFS 5.6 low, 10% intermediate risk, 18.2% high. Adjuvant! Intermediate: 13.4, 13.9, 43.2. Adjuvant! High: 5, 23.4, 31.5%	CT interaction P=0.031 DRFS, P=0.011 OS, P=0.082 DFS,	Concordance with Adjuvant! 0.49,Adjuvant! OS interaction CT p=0.009, DRFS p=0.219
Freitas & Simon[71]	2011	D	N=22, ER+, HER- early stage	Low 11, Inter/high 11			Kappa Adjuvant! 0.091, Adjuvant! Transbig 0.182, NCCN 0.091
Aktas[72]	2013	D	N=68, HER2-	Low 30 Intermediate 29 High 9			Correlation RS – PR p=0.006, with G3 p=0.002, low Ki67 p<0.001
Acs[73]	2013	D	N=106, low grade, ER+	Low 68 Intermediate 38 High 0	Comparison with mammostrat (immunohistochemistry)		RS and NPI r=-0.0737, p=0.4527
Kok[74]	2009	D	N=246 , M+, TAM treated	Low-intermediate 28 High 41	78-gene TAM response, Oncotype DX and HOXB13-IL17BRratio-TTP:HR 2.2 (1.3–3.7, P=0.005),2.3 (1.3–4.0, P=0.003) & 4.2 (1.4–12.3, P =0.009)	Multivariable model 1.94 (1.01-3.73); p=0.048	Concordance 45-61%
PAM50 / Prosigna							
Martin[75]	2013	B	N=820		OS low PAM50 HR 0.23 (0.09-0.57), p<0.001	Interaction PAM50-treatment: p=0.006 cont. p=0.019 groups	
Liu[76]	2015	B	N=1094, <=60 yrs, N+/high risk No	Low 3.4% Intermediate 17.9%, high 78.7%	Higher ROR worse RFS p=0.03. Multivariable ROR high vs low/int HR 1.98 (0.53-7.45; p=0.311)	Subtypes p=0.002 multivariable model. Not predictive treatment effect p-interaction=0.23	
Sestak[77]	2015	B	N=2137, HR+, postmenopausal		DR High risk: 16.6 (13.1-20.9), intermediate 8.3 (6.1-11.2), low 2.4 (1.6-3.5). HR high 6.9 (4.54-10.47), intermediate HR 3.26 (2.07-5.13) compared to low	Added to clinical factors: Univariable LRchi2=67.94; Multivariable LRchi2=35.25	*r*=0.36

Table 2 continued

Table 2 continued

Authors	Year	LOE	Patients (N)	Low/High risk (n)	Outcome	HR (95%CI)	Concordance
Gnant[78]	2014	B	N=1478, ER+, postmenopausal	Low 502 Intermediate 478 High 498	DR ROR HR 1.03 (1.02–1.04, P < 0.0001);log-likelihood test: ΔLRχ2 = 53.49; P < 0.0001 DRFS 10-yrs low risk 96.7 (94.6–98.0, intermediate 91.3% (88.1–93.8), high 79.9% (75.7–83.4)		Spearman's correlation coefficient: 0.32, P < 0.0001
Filipits[79]	2014	B	N=1246	Low 460 Intermediate 416 High 370	Late DRFS compared to clinical factors : DLRc2 15.32, P < 0.001. 15-yrs DRFS low 97.6 (94.7-98.9), intermediate 90.9 (85.9-94.2), high 82.5 (74.8-88.1)		
EndoPredict							
Fitzal[80]	2015	B	N=1324, ER+, HER2-, Postmenopausal	High: 683 Low: 641	10-year LRFS: 91% 10-year LRFS: 97.5%	HR 1.31 (1.16-1.48); p<0.005	
Dubsky[81]	2013	B	N=1702, ER+, HER2-, early stage, Postmenopausal	High 37% Low 63%	Low: 10-year DM 95.3% (93.4-97.3)		58-61% high/int according to clinical guidelines to low risk
Dubsky[82]	2013 (BJC)	B	N=1702, ER+, HER2-, Postmenopausal	High 51% Low 49%	Low: 1.8% DM 10 yrs	0-5 yrs: 1.20 (1.10-1.31) >5 yrs: 1.28 (1.10-1.48)	C index with clinic-pathological parameters: 0.716
Filipits[11]	2011	B	N=378+1324		DR EPclin recurrence rates 4% and 4% in EPclin low-risk; 28% and 22% in EPclin high-risk (P < 0.001) and ABCSG-8 (P < 0.001), respectively.	Multivariable model 1.19 (1.04-1.36) and 1.26 (1.15-1.38)	

LOE = level of evidence according to Simon and Hayes[83], LN- = Lymph Node negative, LN+ = Lymph Node positive, LRFS = local recurrence free survival, BC = breast cancer, DM = distant metastases, CR = clinical risk, GR = Genomic Risk, TTP = Time to Tumour Progression, LRR = loco-regional recurrence, BCI = breast cancer index assay, TTP = time to progression, NPI = Nottingham prognostic index

Clinical validation

A total of 50 studies was identified assessing the clinical benefit of the genomic assays; 21 assessing the MammaPrint, 20 assessing the OncotypeDX, 5 assessing the PAM50/Prosigna and 4 assessing the Endopredict. Most of the studies were retrospectively stratifying the cohort in separate risk categories determined by the test, and showing a difference in either distant metastasis-free, disease-free or overall survival. Table 2 shows the results of the retrospective included studies, according to test and patient inclusion. In general, the studies are difficult to compare due to different patient inclusion and outcome measures. All published studies showed a good differentiation in high and low risk and were associated with survival (both Distant Metastases/Recurrence Free Survival (DMFS/DRFS) as Overall Survival (OS)). In more detail, MammaPrint was reported to be of significant prognostic value for patients with lymph node negative breast cancer and the results of the test correlated well with Adjuvant!, St Gallen and NIH guidelines and the NPI. For lymph node positive disease, the hazard ratios for DMFS and Breast Cancer Specific Survival (BCSS) showed a significant difference in prognosis for low versus high risk according to MammaPrint. In the remaining articles (without specific classification or LN- and LN+ combined) the MammaPrint was also of prognostic value; most of the results showed a significant difference in outcome between low and high risk.

With respect to OncotypeDX, most of the studies in patients with LN negative disease studied the DRFS and showed a significant difference in outcome between low, intermediate and high risk patients. Paik et al showed a statistical different effect of chemotherapy in the three risk groups with a significant interaction term between chemotherapy and the Recurrence Score. One case-control study showed a significant difference between both groups. Besides, the study in LN+ disease also showed a significant interaction between the RS and clinical benefit of chemotherapy for the first 5 years after treatment. The remaining studies (combined LN- and LN+ and one study in patients with metastatic disease) showed a good discrimination between the three risk groups and a significant difference in outcome in most of the studies.

Studies that used the PAM50 showed a good discrimination, and a significant interaction between treatment and outcome in one study, this was however not confirmed in Liu et al. Three studies showed a significant association with distant recurrences. For studies that used EndoPredict differences between high and low risk

were associated with outcome or showed a low proportion of distant metastases in the low risk group.

Both the PAM50/Prosigna and EndoPredict have a quality B level of evidence in all of their validation studies by performing them in established clinical trials, according to Simon et al.[83] For MammaPrint one level A trial is available[22], all other studies are level C quality or lower. For OncotypeDX, there is a mix of two level A trials[8,36], some level B studies showing predictive capacities of OncotypeDX, and level C/D studies in retrospective or case-control studies. All level A evidence will be discussed in the next paragraphs.

MINDACT

The MINDACT trial evaluated the use of the MammaPrint together with Adjuvant Online, an online tool using clinicopathological information for risk stratification.[22] Patients with discordant risks based on the clinical and genomic assessment, were randomized between chemotherapy or no chemotherapy. The primary study subgroup were the patients with a clinical high and genomic low risk tumour who were randomly allocated to receive no chemotherapy. The distant metastasis-free survival of this group was 94.7% at 5 years, which was significantly higher compared to a pre-determined null hypothesis of 92%. Therefore, it was concluded that the prognosis of these clinically high-risk, but genomic low risk patients without chemotherapy was good enough to justify the abstention of chemotherapy.

The trial is labelled as phase 3 RCT and the results are regarded as level IA evidence. However, the design of the primary analysis is that of a cohort study, since it only assessed the patients who had a discordant risk and did not receive chemotherapy. In a secondary per-protocol analysis, comparing the c-high/g-low patients with and without chemotherapy, a HR of around 0.65 was shown in favour of chemotherapy, which was significant for DFS (90.3% vs 93.3%, p=0.026), but not for DMFS (94.8 vs 96.7, p=0.106) or OS (97.3 vs 98.8, p=0.245). In summary, although the prognosis of this clinically high-risk group is good without chemotherapy, it is significantly better when receiving chemotherapy.

Another secondary outcome is the effect of chemotherapy in patients who were clinically assessed as low-risk, but with a genomic high risk profile. In this subgroup, no statistically significant benefit of chemotherapy was observed for either DMFS (HR

0.90 95% CI 0.40-2.01), DFS (HR 0.74 95% CI 0.40-1.39) or OS (HR 0.72, 95% CI 0.23-2.24), indicating that a high-risk MammaPrint test result does not predict an effect of chemotherapy for these low-risk patients. Although this analysis is underpowered, and no formal interaction test was performed, the authors conclude that the MammaPrint failed to show its value as a predictive biomarker, not being capable of identifying patient who would benefit from chemotherapy.

TAILORx

The TAILORx trial was designed to assess the clinical use of OncotypeDX to decide on the chemotherapy administration, especially in the intermediate risk group. For this, 10,273 patients were enrolled, who all had ER- and/or PR-positive, node-negative disease but did have an indication for chemotherapy based on the NCCN-guidelines. Low-risk patients (based on Recurrence Score) received endocrine therapy only; high risk patients received both endocrine and chemotherapy. Intermediate risk-group patients were randomly allocated to either endocrine therapy alone or a combination of endocrine and chemotherapy. Until now, only the results of the low-risk patients were published.[8]

A total number of 1626 patients with a low-risk OncotypeDX test received no chemotherapy. The rate of DFS at 5 years was 93.8%, the freedom from distant recurrence was 99.3% and the overall survival was 98%. Similar to the MINDACT trial, this shows that genomic testing can identify patients with a good prognosis without chemotherapy, despite a clinical indication for chemotherapy.

In a similarly designed trial (RxPonder), node-positive patients with HR+ breast cancer and a low or intermediate test result are randomly assigned to hormone therapy with or without chemotherapy.[84] Results of this trial will show whether it is safe to withhold chemotherapy based on a low or intermediate test result population despite the high-risk nodal status.

WSG PlanB

In the West German Study Group Phase III PlanB Trial, 3198 clinically high-risk patients were enrolled, including 41.1% with node-positive disease. Although originally designed to compare two regimes of chemotherapy, after inclusion of 274 patients the study was amended to omit chemotherapy in patients with a low-risk OncotypeDX test result, despite their high clinical risk.[36]

In this high-risk population, 348 patients received no chemotherapy based on a low-risk Recurrence Score of <12. At 3 years of follow-up, the disease-free survival was 98.4% in this subgroup, indicating again that genomic subtyping can identify a clinically high-risk subgroup with an excellent prognosis without chemotherapy, although longer follow-up is warranted for definite conclusions. Similar to the TAILORx, this study used an alternative cut-off for low-risk scores, which needs to be considered when interpreting the results.

Clinical Utility

A total of 28 studies which evaluated the clinical utility of assays has been identified, of which 22 for OncotypeDX, four for MammaPrint, and one for both Prosigna and Endopredict. Almost all studies compared the (hypothetical) application of chemotherapy for the same patient, with and without the results of the genomic test. In general, de-escalation from chemotherapy to no therapy or endocrine therapy alone was higher than the escalation towards chemotherapy, which led to a decrease in chemotherapy use for all tests. When the results were pooled per assay, the decrease in chemotherapy was the most pronounced for OncotypeDX (45.7% from chemotherapy to endocrine therapy alone or no adjuvant therapy) compared to MammaPrint (32.2% decrease) (table 3). However, these pooled results should be interpreted carefully, since there is a large difference in the number of studies per test, the baseline patient populations and study designs.

For OncotypeDX, three other studies evaluated the use of chemotherapy in population studies.[113–115] Two of them observed a decrease in chemotherapy use during the designated years, and an increase in genomic testing.[114,115] However, no direct relation was observed between both results. In the study of Su et al, performed in a US medicare population between 2008 and 2011, no difference in the use of chemotherapy was observed despite an increase of assay use from 9 to 17.2%.[113]

Table 3 Clinical utility, according to test and nodal status

Clinical utility

Authors	Year	Patients (N)	% chemotherapy before test	% chemotherapy after test	% change to chemotherapy	% change to HT/no therapy
MammaPrint						
Lymph node negative						
Drukker[85]	2014	N=414, T1-3	49	37	4.3	29.1
Early stage breast cancer (combined LN- and LN+, other groups or not specified)						
Pohl[86]	2016	N=107, HR+HER2-	56.1	39.2	40	62
Exner[87]	2014	N=75, grade 1 or 2, T 1-3cm, HR+HER2-	41.3	33.3	9.1	32.3
Cusumano[88]	2014	N=194, T1-3N0-1	60.8	60.8	34.6	22.3
Subtotal MammaPrint						
		N=790	52.1	42.8	17.0	32.2
OncotypeDX						
Lymph node negative						
Ozmen[89]	2016	N=165, T1-3N0-1mic, HR+HER2-	55.8	37	13.7	44.6
Levine[90]	2016	N=972, T1-4N0-1mic, HR+HER2-	22*	20.7	10.9	62.6
Leung[91]	2016	N=146, T1-3N0-1mic, HR+	52.1	37.7	4.3	31.6
Gligorov[92]	2015	N=100, T1-3N0-1mic, HR+HER2-	52	25	10.9	61.2
Lee[93]	2015	N=212, T1-3N0-1mic, HR+	70.7	22.1	9.7	72.7
Jaafar[94]	2014	N=47, T1-2N0, HR+HER2-	48.9	25.5	4.2	52.2
Davidson[95]	2013	N=150, T1-3N0, HR+HER2-	41.3	31.3	17	48.4
Holt[96]	2013	N=142, T1-3N0-1mic, HR+	40.1	30.3	14.1	45.6
Biroschak[97]	2013	N=50, T1-3N0, HR+	72	70	28.6	13.9
Ademuyiwa[98]	2011	N=276, T1-3N0 HR+HER2-	45.3	32	22.5	56.8
Albanell[99]	2011	N= 107, T1-3N0, ER+HER2-	37	27	17.6	56.4
Lo[100]	2010	N=89, T1-2N0, HR+	47.2	25.9	6.5	47.6
Henry[101]	2009	N=29, T1-3N0, HR+	45	28	13	54
Oratz[102]	2007	N=74, T1-3N0, HR+	48	48	20	21.2
Early stage breast cancer (combined LN- and LN+, other groups or not specified)						
Kuchel[103]	2016	N=137, T1-3N0-1, HR+HER2-	50.4	27.7	18.2	62.3
Bargallo[104]	2014	N=96 , T1-3N0-1 ER+HER2-	48	31	16	45.7
Yamauchi[105]	2014	N=124, T1-3N0-1, HR+HER2-	51	24	11.5	63.5

Table 3 continued

Authors	Year	Patients (N)	% chemotherapy before test	% chemotherapy after test	% change to chemotherapy	% change to HT/no therapy
Fried[106]	2014	N=111, T1-3N0-1, HR+	29.7	27.9	14.1	39.4
Cheung[107]	2014	N=64,T1-2N0-1, HR+HER2-	61	55	16	20.5
Eiermann[108]	2013	N=366, T1-3N0-1, HR+HER2-	57	46	25	38
De Boer[109]	2013	N=151, T1-3N0-1, HR+HER2-	44.4	37.1	15.5	35.8
Geffen[110]	2011	N=135, T1-2N0-1	47	36	13.9	38.1
Subtotal OncotypeDX						
		N=3743	50.2	30.6	14.6	51.1
PAM50 / Prosigna						
Martin[111]	2015	N=200, T1-2N0, HR+HER2-	30%	28%	12.9%	37.3%
EndoPredict						
Muller[112]	2013	N=167, T1-3N1-3, HR+HER2-	63.8%	47.7%	34%	53.2%

*not included in pooled data, since pre-test chemotherapy also included 34% unsure

Two other studies evaluated the use of chemotherapy between patients with and without genomic testing.[116,117] In the large study performed by Ray et al (n=7004), 22% of chemotherapy was observed in patients without testing, whereas 26% used chemotherapy after genomic profiling. In contrast, Stemmer et al (n=951) observed in a node-positive population, a 70% chemotherapy use without testing and a 24.5% chemotherapy use after genomic testing.

In a similar study design, Kuijer et al observed a 10% lower rate of chemotherapy for patients with genomic testing using MammaPrint.[118]

Economic value

Forty-four original economic evaluations were found, of which 32 on Oncotype DX, 7 on MammaPrint, 1 on EndoPredict and 4 direct comparisons between tests (Table 4). Most evaluations compared genomic testing to a variety of strategies without genomic testing; four evaluations were head-to-head comparisons between genomic policies. Of the evaluations, 5 only estimated costs (CMAs), 1 estimated life years without QALYs (CEA) and 38 estimated QALYs (CUAs).

Methodologically, only 2 evaluations (both CMA) compared measured outcomes between two actual patient groups with and without genomic testing.[113,119] The remaining 42 evaluations all used mathematical (mostly Markov) modelling to compare estimated outcomes for different policies, for the same actual or hypothetical group of patients. These mathematical models typically estimated a decrease in chemotherapy (because the shift to low risk exceeds the shift to high risk), a decrease in recurrence (because the decrease in high risk exceeds the increase in low risk), and an increase in life years and QALYs (due to the decrease in recurrence and toxicity). Total health care costs may go up or down, depending on the balance between the assay costs and savings on chemotherapy and recurrence. Three studies also included savings on productivity.[120-122]

Figure 2 shows the estimated impact of genomic testing on QALYs and costs, according to the 40 evaluations comparing genomic testing to a strategy without genomic testing. The horizontal axis shows the impact on QALYs: all studies but one[123] reported that genomic testing resulted in better patient outcome with a positive impact on QALYs. The vertical axis shows the impact on costs: genomic testing was cost saving in 14 (35%) evaluations and cost increasing in 26 (65%) of the evaluations. On average, total costs increased by 449 euro per patient with an improvement on patient outcome of 0.16 life years and 0.20 QALYs. In general, there were no apparent differences between the estimated outcomes for the different genomic tests. Also, the range of costs was comparable in node-negative and node-positive patients, but the estimated QALY gain was larger in node-negative patients (on average, 0.24 versus 0.07 QALYs). Considering the improvement in patient outcome, genomic testing was cost-effective in 36 (90%) of the evaluations, i.e. below the dashed 40,000 euro-per-QALY line.

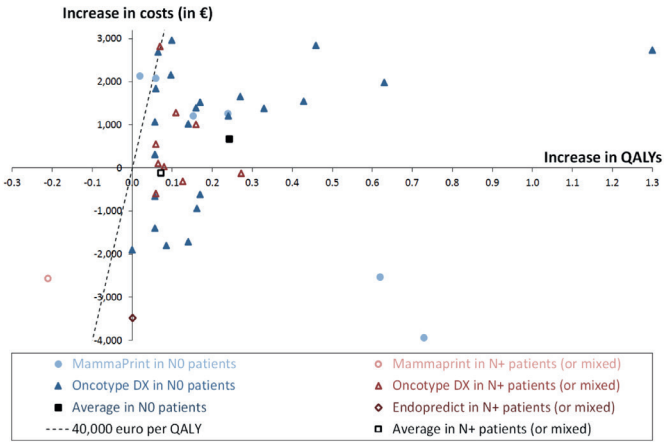


Figure 2. Estimated impact on costs and quality-adjusted life years (QALYs) per economic evaluation, according to test and nodal status

Table 4 Economic evaluations, according to test and nodal status

Authors	Year	Comparator	Patient group	Country	Impact on costs	Impact on QALYs	Impact on life years	Economic conclusion
MammaPrint compared to no genomic testing								
Lymph node negative								
Bonastre[124]	2014	Adjuvant! Online	No	France	€ 2037	0.02	0.01	€ 134,000 per QALY
Chen[125]	2010	Adjuvant! Online	No	US	$ 1440	0.153	0.143	$ 10,000 per QALY
Exner[87]	2014	Usual care	No HR+ HER2-	NL	€ -3779	0.73	-	Dominant
Kondo[126]	2012	Best practice	No ER+ HER2-	Japan	$ 2571	0.06	0.048	$ 43,044 per QALY
Retèl[127]	2010	Adjuvant! Online	No ER+	NL	€ 1130	0.24	0.2	€ 4,614 per QALY
Retèl[128]	2013	Adjuvant! Online	No ER+	NL	€ -2401	0.62	-	Dominant
Lymph node positive (or mixed)								
Oestreicher[123]	2005	Best practice	N≥0 stage≤II pre-menopausal	US	$ -2882	-0.21	-	$ 13,724 per QALY (in favor of BP)
OncotypeDX compared to no genomic testing								
Lymph node negative								
Bacchi[129]	2010	Usual care	No ER+	Brazil	$ -794	-	-	Cost saving
Cosler[130]	2009	Chemotherapy+ Tamoxifen	No ER+	US	$ -2256	0	-	Dominant
Davidson[95]	2013	Usual care	No ER+ HER2-	Canada	CAN$ 2188	0.33	0.31	CAN$ 6,630 per QALY
Epstein[119]	2015	Usual care	No ER+	US	$ 1367	-	-	Cost increasing
Hannouf[131]	2012	Usual care	No HR+	Canada	CAN$ 2879	0.059	-	CAN$ 48,493 per QALY
Holt[96]	2013	Usual care	No-1 ER+	UK	£ 888	0.14	0.16	£ 6,232 per QALY
Hornberger[132]	2005	Usual care	No ER+	US	$ -1160	0.162	-	Dominant
Hornberger[133]	2011	Best practice	No ER+	US	$ -2028	0.086	-0.0421	Dominant
Jahn[134]	2015	Adjuvant! Online	No HR+ HER2-	Austria	€ 2750	0.46	0.59	€ 5,978 per QALY
Katz[120]	2015	Usual care	No HR+ HER2-	France	€ -602	0.17	0.18	Dominant
Klang[135]	2010	Usual care	No ER+	Israel	$ 1828	0.17	-	$ 10,770 per QALY

Table 4 continued

Authors	Year	Comparator	Patient group	Country	Impact on costs	Impact on QALYs	Impact on life years	Economic conclusion
Kondo[136]	2008	Best practice	No HR+	Japan	$ 2516	0.097	0.083	$ 30,137 per QALY
Kondo[137]	2011	Best practice	No ER+	Japan	$ 2407	0.63	-	$ 3,848 per QALY
Lamond[138]	2012	Usual care	No ER+	Canada	CAN$ 2585	0.27	-	CAN$ 9,591 per QALY
OHTA[139]	2010	Adjuvant! Online	No HR+ HER2-	Ontario	CAN$ 4168	1.3	-	CAN$ 3,206 per QALY
Paulden[140]	2013	Adjuvant! Online	No HR+ HER2-	Canada	CAN$ 2460	0.429	0.53	CAN$ 5,734 per QALY
Reed[121]	2013	Adjuvant! Online	No ER+	US	$ 1741	0.16	0.19	$ 10,788 per QALY
Smyth[141]	2015	Best practice	No ER+	Ireland	€ -1361	-	-	Cost saving
Su[113]	2016	Usual care	No HR+ HER2-	US	$ 400	-	-	Cost increasing
Tsoi[142]	2010	Adjuvant! Online	No HR+	Canada	CAN$ 4102	0.065	0.064	CAN$ 63,064 per QALY
Vataire[122]	2012	Usual care	No ER+ HER2-	France	€ -1600	0.14	0.15	Dominant
Ward[143]	2013	Usual care	No ER+ HER2-	UK	£ 2575	0.1	-	£ 29,502 per QALY
Yamauchi[144]	2014	Usual care	No ER+	Japan	$ 1536	0.241	-	$ 6,368 per QALY
Lymph node positive (or mixed)								
Bargalló-Rocha[145]	2015	Usual care	N3 HR+ HER2-	Mexico	$ 129	-	0.068	$ 1,914 per LY
Blohmer[146]	2013	Usual care	N3 ER+ HER2-	Germany	€ -561	0.06	0.06	Dominant
Hall[147]	2012	Chemotherapy	N+ ER+	UK	£ 860	0.16	0.15	£ 5,529 per QALY
Hannouf[148]	2014	Usual care	N+ HR+ post-menopausal	Canada	CAN$ 36.2	0.08	-	CAN$ 464 per QALY
Kip[149]	2015	Usual care	N1 ER+	NL	€ 1236	0.11	-	€ 11,236 per QALY
Kondo[137]	2011	Best practice	N+ ER+	Japan	$ 3434	0.07	-	$ 49,059 per QALY
Lamond[138]	2012	Usual care	N+ ER+	Canada	CAN$ 864	0.06	-	CAN$ 14,844 per QALY
Nerich[150]	2014	Usual care	N1 ER+ HER2-	France	€ -128	-	-	Cost saving
Vanderlaan[151]	2011	Best practice	N+ ER+ HER2-	US	$ -384	0.127	-	Dominant

Table 4 continued

Authors	Year	Comparator	Patient group	Country	Impact on costs	Impact on QALYs	Impact on life years	Economic conclusion
EndoPredict compared to no genomic testing								
Lymph node positive (or mixed)								
Blank[152]	2015	Best practice	N≥0 ER+ HER2-	Germany	€ -3388	0.002	-0.037	Dominant
Head-to-head comparisons								
Mislick[153]	2014	Mammostrat vs OncotypeDX	No ER+	US	$ -2268	-0.005	-0.002	$ 453,600 per QALY (in favor of Mammostrat)
Retèl[154]	2012	MammaPrint vs OncotypeDX	No ER+	NL	€ -1475	0.08	-0.14	Mammaprint dominant
Seguí[155]	2014	MammaPrint vs OncotypeDX	No ER+ HER2-	Spain	€ 1085	0.745	0.863	€ 1,457 per QALY (in favor of Mammaprint)
Yang[156]	2012	MammaPrint vs OncotypeDX	No ER+	US	$ -6284	0.097	-	Mammaprint dominant

Discussion

In this systematic review, we evaluated four commercially available prognostic genomic profiles on four selected crucial aspects. On all aspects, the tests are well-studied, with multiple well-designed and well-performed studies available. It is apparent that on the level of quantity, MammaPrint and especially OncotypeDX are more extensively studied compared to the more recently developed Endopredict and Prosigna/PAM50 assay. At this time of development, both OncotypeDX and MammaPrint are suitable assays which can be helpful in the clinical setting. However, this review also identified some caveats which will need to be addressed before genomic profiling can be optimally applied.

Assay development and methodology

The first topic for improvement is the identification of a subgroup that benefits most from genomic profiling. This has already been investigated for OncotypeDX, and to a lesser extent for MammaPrint. For Prosigna and Endopredict we did not identify

studies that studied for which clinicopathological subtypes genomic profiling is valuable. In general, the studies show that patients with grade 3, PR- and a high Ki-67 have no benefit from testing, since they are almost always high-risk. In contrast, patients with grade 1, ER+PR+ and Ki-67 <10% have no benefit from testing either, since (almost) all of them had a low-risk result. As suggested by the flowchart build by Allison et al, all other patients would have an indication for genomic profiling.[38] However, most of these studies were performed in a node-negative cohort. MINDACT has shown that despite node-positive disease, it could be considered to withhold chemotherapy at a low genomic risk score. Therefore, it is crucial that this test-result predicting model is validated and adjusted in large trial cohorts like MINDACT and the WSG Plan-B trial.

Clinical validation

One of the most important (theoretical) benefits of a genomic profiling test is the selection of patients in which the treatment with adjuvant chemotherapy will have a significant benefit. Currently, this task of genomic profiles is mainly performed by their prognostic capacities; i.e. the ability to identify patients with a poor prognosis for recurrence or survival. However, the results of the studies in this review, especially that of MINDACT, show that this does not automatically translate into a benefit of chemotherapy for these higher-risk patients. So far, no genomic test has shown it's predictive capacities in a prospective trial design. The only evidence for a predictive value was obtained in two prospective studies conducted on archived tissue (prospective-retrospective design) in which the OncotypeDX retrospectively identified patients that benefit more from chemotherapy to which they were randomly allocated.[62,65]

Clinical utility

Currently, the clinical consensus on adjuvant chemotherapy is that we are most likely over-treating our patients, since we are not capable of identifying patients that will or will not benefit from chemotherapy using the current clinicopathological parameters.[157,158] It is no surprise that the studies evaluating the clinical utility of genomic profiling especially show a reduction in chemotherapy use. However, absolute numbers should be interpreted carefully, since some tests are less frequently studied than others, which increases the risk of bias and skewed data. Interestingly, in retrospective population-based cohorts, implementation of genomic testing did not lead to a reduction in chemotherapy use.[113-115] This is in accordance with Petkov et al, who retrospectively matched OncotypeDX use with SEER registry data for

over 40,000 patients.[159] Although the risk categories were indeed prognostic for five-year breast-cancer-specific mortality in this real-life population, patients with node negative, HR+, HER2- breast cancer which underwent testing (n=40,134, 22.7% chemotherapy) had no lower chemotherapy use compared to patients that were not tested (n=144,056, 22.2% chemotherapy). Therefore, conclusions about genomic profiling leading to decrease in chemotherapy cannot be drawn from these analyses.

Economic value

Our review of economic evaluations identified 44 original publications, where earlier reviews included at most 11 or 18 published evaluations.[160,161] Except for the oldest evaluation[123], all studies reported improved patient outcome in terms of QALYs. Despite estimated savings on chemotherapy, recurrence and productivity, a small majority (65%) of the evaluations estimated that genomic testing resulted in an increase in total costs. Nevertheless, most evaluations (90%) estimated that genomic testing is cost-effective, with costs that are acceptable in relation to patient outcome. These economic results should be considered with caution. Firstly, the separate evaluations should not be interpreted as independent primary studies, because the models obtain their data from overlapping sources: mostly the diagnostic data are taken from the landmark trials and then applied to the care patterns of a particular country. Secondly, the economic studies generally evaluate the use of genomic testing in large groups of women, instead of trying to combine genomic profiling with other prognostic factors to identify those individual women for whom genomic testing does not have sufficient added value or could even be harmful. And thirdly, compared to trials, economic evaluations are more likely to suffer from publication bias.

Future perspectives

In the near future, trial results from RxPonder, TAILORx and WSG plan-B will become available, contributing to understanding the role of OncotypeDX in daily practice in both node-positive and node-negative disease. Furthermore, subgroup analyses and long-term follow-up of MINDACT will follow later and help define the place for MammaPrint in the diagnostic process, and the long-term safety of withholding chemotherapy in high-risk patients, based on a low-risk test result. The OPTIMA trial, randomizing high-risk ER+HER2- patients between standard chemotherapy, or treatment directed by Prosigna test-results will be the first trial to show level A evidence for the Prosigna/PAM50 test.

Another interesting development is the use of gene expression assays for the indication of endocrine therapy. Very recently, a retrospective analysis from Sweden identified an ultra-low category within the low-risk category of MammaPrint (15% of all patients, 26% of low-risk patients).[162] Patients with this ultra-low risk score (n=98) had a breast cancer specific survival of 94% at 20 years without any adjuvant therapy, and 97% at 20 years with just 2 years of tamoxifen, whereas 5+ years of therapy is the current standard for these patients.[163] Upon validation, these findings could lead to the implementation of gene expression assays in the indication for adjuvant endocrine therapy.

Conclusions

In summary, in this systematic review we have evaluated the four most frequently used assays in Europe on four relevant aspects. Regarding the amount of evidence, there is a clear separation between the more established MammaPrint and OncotypeDX on one hand, and the newer Prosigna and Endopredict on the other hand. Comparing MammaPrint and OncotypeDX, both assays have shown to be a useful prognostic tests which could lead to a reduction in chemotherapy use, with in general a favourable cost-benefit ratio. Both the MammaPrint and OncotypeDX have shown in prospective trials that a patient with a low-risk result can safely forego chemotherapy, despite clinical risk factors. In contrast, the benefit of chemotherapy with a high-risk test result has so far only been shown for OncotypeDX, albeit in retrospective analyses of archived tissue of prospective trials. Therefore, there is still a need for further prospective studies on all evaluated assays.

Acknowledgements

We sincerely thank prof H. Putter PhD, medical statistician in the Leiden University Medical Center, for checking and approving our study design and methods. Furthermore, we thank J.W. Schoones MA, librarian in the Leiden University Medical Center, for assistance in the study design, and designing the search strategy.

References

1. (EBCTCG) EBCTCG. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717.
2. Early Breast Cancer Trialists' Collaborative G. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. Lancet 2012;379:432-44.
3. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. The Lancet;378:1812-23.
4. Ballman KV. Biomarker: Predictive or Prognostic? Journal of Clinical Oncology 2015;33:3968-71.
5. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530-6.
6. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999-2009.
7. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. New England Journal of Medicine 2004;351:2817-26.
8. Sparano JA, Gray RJ, Makower DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. N Engl J Med 2015;373:2005-14.
9. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160-7.
10. Nielsen T, Wallden B, Schaper C, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. BMC Cancer 2014;14:177.
11. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res 2011;17:6012-20.
12. Glas AM, Floore A, Delahaye LJMJ, et al. Converting a breast cancer microarray signature into a high-throughput diagnostic test. Bmc Genomics 2006;7:278.
13. Ach RA, Floore A, Curry B, et al. Robust interlaboratory reproducibility of a gene expression signature measurement consistent with the needs of a new generation of diagnostic tools. BMC Genomics 2007;8:148.
14. Mittempergher L, de Ronde JJ, Nieuwland M, et al. Gene expression profiles from formalin fixed paraffin embedded breast cancer tissue are largely comparable to fresh frozen matched tissue. PLoS One 2011;6:e17163.
15. Sapino A, Roepman P, Linn SC, et al. MammaPrint molecular diagnostics on formalin-fixed, paraffin-embedded tissue. J Mol Diagn 2014;16:190-7.
16. Beumer I, Witteveen A, Delahaye L, et al. Equivalence of MammaPrint array types in clinical trials and diagnostics. Breast Cancer Res Treat 2016;156:279-87.
17. Gyanchandani R, Lin Y, Lin HM, et al. Intratumor Heterogeneity Affects Gene Expression Profile Test Prognostic Risk Stratification in Early Breast Cancer. Clin Cancer Res 2016;22:5362-9.
18. Drury S, Salter J, Baehner FL, et al. Feasibility of using tissue microarray cores of paraffin-embedded breast cancer tissue for measurement of gene expression: a proof-of-concept study. Journal of Clinical Pathology 2010;63:513-7.
19. Müller BM, Brase JC, Haufe F, et al. Comparison of the RNA-based EndoPredict multigene test between core biopsies and corresponding surgical breast cancer sections. Journal of Clinical Pathology 2012;65:660-2.
20. Denkert C, Kronenwett R, Schlake W, et al. Decentral gene expression analysis for ER+/Her2- breast cancer: results of a proficiency testing program for the EndoPredict assay. Virchows Arch 2012;460:251-9.
21. Kronenwett R, Bohmann K, Prinzler J, et al. Decentral gene expression analysis: analytical validation of the Endopredict genomic multianalyte breast cancer prognosis test. BMC Cancer 2012;12:456.
22. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med 2016;375:717-29.
23. Elloumi F, Hu Z, Li Y, et al. Systematic Bias in Genomic Classification Due to Contaminating Non-neoplastic Tissue in Breast Tumor Samples. BMC Medical Genomics 2011;4:54-.

24. Stein RC, Dunn JA, Bartlett JM, et al. OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. Health technology assessment (Winchester, England) 2016;20:xxiii-xxix, 1-201.
25. Bartlett JMS, Stein RC, Bayani J, et al. Comparison of multiparameter tests in the UK OPTIMA-Prelim trial. Cancer Research 2015;75.
26. Alvarado MD, Prasad C, Rothney M, et al. A Prospective Comparison of the 21-Gene Recurrence Score and the PAM50-Based Prosigna in Estrogen Receptor-Positive Early-Stage Breast Cancer. Adv Ther 2015;32:1237-47.
27. Varga Z, Sinn P, Fritzsche F, et al. Comparison of EndoPredict and Oncotype DX test results in hormone receptor positive invasive breast cancer. PLoS One 2013;8:e58483.
28. Gevensleben H, Gohring UJ, Buttner R, et al. Comparison of MammaPrint and TargetPrint results with clinical parameters in German patients with early stage breast cancer. Int J Mol Med 2010;26:837-43.
29. Chaudhary LN, Jawa Z, Szabo A, et al. Relevance of progesterone receptor immunohistochemical staining to Oncotype DX recurrence score. Hematol Oncol Stem Cell Ther 2016;9:48-54.
30. Dialani V, Gaur S, Mehta TS, et al. Prediction of Low versus High Recurrence Scores in Estrogen Receptor-Positive, Lymph Node-Negative Invasive Breast Cancer on the Basis of Radiologic-Pathologic Features: Comparison with Oncotype DX Test Recurrence Scores. Radiology 2016;280:370-8.
31. Ingoldsby H, Webber M, Wall D, et al. Prediction of Oncotype DX and TAILORx risk categories using histopathological and immunohistochemical markers by classification and regression tree (CART) analysis. Breast 2013;22:879-86.
32. Sahebjam S, Aloyz R, Pilavdzic D, et al. Ki 67 is a major, but not the sole determinant of Oncotype Dx recurrence score. British journal of cancer 2011;105:1342-5.
33. Auerbach J, Kim M, Fineberg S. Can Features Evaluated in the Routine Pathologic Assessment of Lymph Node-Negative Estrogen Receptor-Positive Stage I or II Invasive Breast Cancer Be Used to Predict the Oncotype DX Recurrence Score? Arch Pathol Lab Med 2010;134:1697-701.
34. Flanagan MB, Dabbs DJ, Brufsky AM, et al. Histopathologic variables predict Oncotype DX recurrence score. Mod Pathol 2008;21:1255-61.
35. Wolf I, Ben-Baruch N, Shapira-Frommer R, et al. Association between standard clinical and pathologic characteristics and the 21-gene recurrence score in breast cancer patients: a population-based study. 2008;112:731-6.
36. Gluz O, Nitz UA, Christgen M, et al. West German Study Group Phase III PlanB Trial: First Prospective Outcome Data for the 21-Gene Recurrence Score Assay and Concordance of Prognostic Markers by Central and Local Pathology Assessment. J Clin Oncol 2016;34:2341-9.
37. Bradshaw SH, Pidutti D, Gravel DH, et al. Predicting OncoDx recurrence scores with immunohistochemical markers. Applied immunohistochemistry & molecular morphology : AIMM 2013;21:490-6.
38. Allison KH, Kandalaft PL, Sitlani CM, et al. Routine pathologic parameters can predict Oncotype DX recurrence scores in subsets of ER positive patients: who does not always need testing? 2012;131:413-24.
39. Williams DJ, Cohen C, Darrow M, et al. Proliferation (Ki-67 and phosphohistone H3) and oncotype DX recurrence score in estrogen receptor-positive breast cancer. Applied immunohistochemistry & molecular morphology : AIMM 2011;19:431-6.
40. Wittner BS, Sgroi DC, Ryan PD, et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. Clin Cancer Res 2008;14:2988-93.
41. Mook S, Schmidt MK, Weigelt B, et al. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. Annals of Oncology 2010;21:717-22.
42. Na KY, Kim KS, Lee JE, et al. The 70-gene prognostic signature for korean breast cancer patients. J Breast Cancer 2011;14:33-8.
43. Drukker CA, Bueno-de-Mesquita JM, Retel VP, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. Int J Cancer 2013;133:929-36.
44. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol 2007;8:1079-87.
45. Bueno-de-Mesquita JM, Linn SC, Keijzer R, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. Breast Cancer Res Treat 2009;117:483-95.

46. Knauer M, Mook S, Rutgers EJ, et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. Breast Cancer Res Treat 2010;120:655-61.
47. Saghatchian M, Mook S, Pruneri G, et al. Additional prognostic value of the 70-gene signature (MammaPrint (R)) among breast cancer patients with 4-9 positive lymph nodes. Breast 2013;22:682-90.
48. Mook S, Schmidt MK, Viale G, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. Breast Cancer Res Treat 2009;116:295-302.
49. Mook S, Knauer M, Bueno-de-Mesquita JM, et al. Metastatic Potential of T1 Breast Cancer can be Predicted by the 70-gene MammaPrint Signature. Annals of Surgical Oncology 2010;17:1406-13.
50. Knauer M, Cardoso F, Wesseling J, et al. Identification of a low-risk subgroup of HER-2-positive breast cancer by the 70-gene prognosis signature. British journal of cancer 2010;103:1788-93.
51. Drukker CA, van Tinteren H, Schmidt MK, et al. Long-term impact of the 70-gene signature on breast cancer outcome. Breast Cancer Res Treat 2014;143:587-92.
52. Drukker CA, Elias SG, Nijenhuis MV, et al. Gene expression profiling to predict the risk of locoregional recurrence in breast cancer: a pooled analysis. Breast Cancer Res Treat 2014;148:599-613.
53. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. Journal of the National Cancer Institute 2006;98:1183-92.
54. Kunz G. Use of a genomic test (MammaPrint (TM)) in daily clinical practice to assist in risk stratification of young breast cancer patients. Arch Gynecol Obstet 2011;283:597-602.
55. Kok M, Koornstra RH, Mook S, et al. Additional value of the 70-gene signature and levels of ER and PR for the prediction of outcome in tamoxifen-treated ER-positive breast cancer. Breast 2012;21:769-78.
56. Ahn SG, Lee HM, Lee HW, et al. Prognostic Discrimination Using a 70-Gene Signature among Patients with Estrogen Receptor-Positive Breast Cancer and an Intermediate 21-Gene Recurrence Score. Int J Mol Sci 2013;14:23685-99.
57. Ishitobi M, Goranova TE, Komoike Y, et al. Clinical utility of the 70-gene MammaPrint profile in a Japanese population. Jpn J Clin Oncol 2010;40:508-12.
58. Toi M, Iwata H, Yamanaka T, et al. Clinical significance of the 21-gene signature (Oncotype DX) in hormone receptor-positive early stage primary breast cancer in the Japanese population. Cancer 2010;116:3112-8.
59. Naoi Y, Kishi K, Tsunashima R, et al. Comparison of efficacy of 95-gene and 21-gene classifier (Oncotype DX) for prediction of recurrence in ER-positive and node-negative breast cancer patients. Breast Cancer Res Treat 2013;140:299-306.
60. Mamounas EP, Tang G, Fisher B, et al. Association Between the 21-Gene Recurrence Score Assay and Risk of Locoregional Recurrence in Node-Negative, Estrogen Receptor-Positive Breast Cancer: Results From NSABP B-14 and NSABP B-20. Journal of Clinical Oncology 2010;28:1677-83.
61. Sgroi DC, Sestak I, Cuzick J, et al. Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population. Lancet Oncol 2013;14:1067-76.
62. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24:3726-34.
63. Tang G, Cuzick J, Costantino JP, et al. Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen receptor-positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. J Clin Oncol 2011;29:4365-72.
64. Yorozuya K, Takeuchi T, Yoshida M, et al. Evaluation of Oncotype DX Recurrence Score as a prognostic factor in Japanese women with estrogen receptor-positive, node-negative primary Stage I or IIA breast cancer. J Cancer Res Clin 2010;136:939-44.
65. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010;11:55-65.
66. Dowsett M, Cuzick J, Wale C, et al. Prediction of Risk of Distant Recurrence Using the 21-Gene Recurrence Score in Node-Negative and Node-Positive Postmenopausal Patients With Breast Cancer Treated With Anastrozole or Tamoxifen: A TransATAC Study. Journal of Clinical Oncology 2010;28:1829-34.
67. Cuzick J, Dowsett M, Pineda S, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. Journal of Clinical Oncology 2011;29:4273-8.

68. Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol 2008;26:4063-71.
69. Le Du F, Gonzalez-Angulo AM, Park M, et al. Effect of 21-Gene RT-PCR Assay on Adjuvant Therapy and Outcomes in Patients With Stage I Breast Cancer. Clinical breast cancer 2015;15:458-66.
70. Tang G, Shak S, Paik S, et al. Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. Breast Cancer Res Treat 2011;127:133-42.
71. Freitas MR, Simon SD. Comparison between Oncotype DX test and standard prognostic criteria in estrogen receptor positive early-stage breast cancer. Einstein (Sao Paulo) 2011;9:354-8.
72. Aktas B, Bankfalvi A, Heubner M, et al. Evaluation and correlation of risk recurrence in early breast cancer assessed by Oncotype DX(R), clinicopathological markers and tumor cell dissemination in the blood and bone marrow. Mol Clin Oncol 2013;1:1049-54.
73. Acs G, Kiluk J, Loftus L, et al. Comparison of Oncotype DX and Mammostrat risk estimations and correlations with histologic tumor features in low-grade, estrogen receptor-positive invasive breast carcinomas. Mod Pathol 2013;26:1451-60.
74. Kok M, Linn SC, Van Laar RK, et al. Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. Breast Cancer Res Treat 2009;113:275-83.
75. Martin M, Prat A, Rodriguez-Lescure A, et al. PAM50 proliferation score as a predictor of weekly paclitaxel benefit in breast cancer. Breast Cancer Res Treat 2013;138:457-66.
76. Liu S, Chapman JA, Burnell MJ, et al. Prognostic and predictive investigation of PAM50 intrinsic subtypes in the NCIC CTG MA.21 phase III chemotherapy trial. Breast Cancer Res Treat 2015;149:439-48.
77. Sestak I, Cuzick J, Dowsett M, et al. Prediction of late distant recurrence after 5 years of endocrine treatment: a combined analysis of patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. J Clin Oncol 2015;33:916-22.
78. Gnant M, Filipits M, Greil R, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Annals of oncology : official journal of the European Society for Medical Oncology 2014;25:339-45.
79. Filipits M, Nielsen TO, Rudas M, et al. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. Clin Cancer Res 2014;20:1298-305.
80. Fitzal F, Filipits M, Rudas M, et al. The genomic expression test EndoPredict is a prognostic tool for identifying risk of local recurrence in postmenopausal endocrine receptor-positive, her2neu-negative breast cancer patients randomised within the prospective ABCSG 8 trial. British journal of cancer 2015;112:1405-10.
81. Dubsky P, Filipits M, Jakesz R, et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. Annals of oncology : official journal of the European Society for Medical Oncology 2013;24:640-7.
82. Dubsky P, Brase JC, Jakesz R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. British journal of cancer 2013;109:2959-64.
83. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. Journal of the National Cancer Institute 2009;101:1446-52.
84. Gonzalez-Angulo AM, Barlow WE, Gralow J, et al. SWOG S1007: A phase III, randomized clinical trial of standard adjuvant endocrine therapy with or without chemotherapy in patients with one to three positive nodes, hormone receptor (HR)-positive, and HER2-negative breast cancer with recurrence score (RS) of 25 or less. Journal of Clinical Oncology 2011;29.
85. Drukker CA, van den Hout HC, Sonke GS, et al. Risk estimations and treatment decisions in early stage breast cancer: Agreement among oncologists and the impact of the 70-gene signature. European Journal of Cancer 2014;50:1045-54.
86. Pohl H, Kotze MJ, Grant KA, et al. Impact of MammaPrint on Clinical Decision-Making in South African Patients with Early-Stage Breast Cancer. The breast journal 2016;22:442-6.
87. Exner R, Bago-Horvath Z, Bartsch R, et al. The multigene signature MammaPrint impacts on multidisciplinary team decisions in ER+, HER2- early breast cancer. British journal of cancer 2014;111:837-42.

88. Cusumano PG, Generali D, Ciruelos E, et al. European inter-institutional impact study of MammaPrint. Breast 2014;23:423-8.

89. Ozmen V, Atasoy A, Gokmen E, et al. Impact of Oncotype DX Recurrence Score on Treatment Decisions: Results of a Prospective Multicenter Study in Turkey. Cureus 2016;8:e522.

90. Levine MN, Julian JA, Bedard PL, et al. Prospective Evaluation of the 21-Gene Recurrence Score Assay for Breast Cancer Decision-Making in Ontario. J Clin Oncol 2016;34:1065-71.

91. Leung RC, Yau TC, Chan MC, et al. The Impact of the Oncotype DX Breast Cancer Assay on Treatment Decisions for Women With Estrogen Receptor-Positive, Node-Negative Breast Carcinoma in Hong Kong. Clinical breast cancer 2016;16:372-8.

92. Gligorov J, Pivot XB, Jacot W, et al. Prospective Clinical Utility Study of the Use of the 21-Gene Assay in Adjuvant Clinical Decision Making in Women With Estrogen Receptor-Positive Early Invasive Breast Cancer: Results From the SWITCH Study. Oncologist 2015;20:873-9.

93. Lee MH, Han W, Lee JE, et al. The clinical impact of 21-gene recurrence score on treatment decisions for patients with hormone receptor-positive early breast cancer in Korea. Cancer Res Treat 2015;47:208-14.

94. Jaafar H, Bashir MA, Taher A, et al. Impact of Oncotype DX testing on adjuvant treatment decisions in patients with early breast cancer: a single-center study in the United Arab Emirates. Asia Pac J Clin Oncol 2014;10:354-60.

95. Davidson JA, Cromwell I, Ellard SL, et al. A prospective clinical utility and pharmacoeconomic study of the impact of the 21-gene Recurrence Score(R) assay in oestrogen receptor positive node negative breast cancer. Eur J Cancer 2013;49:2469-75.

96. Holt S, Bertelli G, Humphreys I, et al. A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the U.K. British journal of cancer 2013;108:2250-8.

97. Biroschak JR, Schwartz GF, Palazzo JP, et al. Impact of Oncotype DX on treatment decisions in ER-positive, node-negative breast cancer with histologic correlation. The breast journal 2013;19:269-75.

98. Ademuyiwa FO, Miller A, O'Connor T, et al. The effects of oncotype DX recurrence scores on chemotherapy utilization in a multi-institutional breast cancer cohort. Breast Cancer Res Treat 2011;126:797-802.

99. Albanell J, Gonzalez A, Ruiz-borrego M, et al. Prospective transGEICAM study of the impact of the 21-gene recurrence score assay and traditional clinicopathological factors on adjuvant clinical decision making in women with estrogen receptor-positive (ER+) node-negative breast cancer. Annals of Oncology 2012;23:625-31.

100. Lo SS, Mumby PB, Norton J, et al. Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol 2010;28:1671-6.

101. Henry LR, Stojadinovic A, Swain SM, et al. The influence of a gene expression profile on breast cancer decisions. J Surg Oncol 2009;99:319-23.

102. Oratz R, Paul D, Cohn AL, et al. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. Journal of Oncology Practice 2007;3:182-6.

103. Kuchel A, Robinson T, Comins C, et al. The impact of the 21-gene assay on adjuvant treatment decisions in oestrogen receptor-positive early breast cancer: a prospective study. British journal of cancer 2016;114:731-6.

104. Bargallo JE, Lara F, Shaw-Dulin R, et al. A study of the impact of the 21-gene breast cancer assay on the use of adjuvant chemotherapy in women with breast cancer in a Mexican public hospital. J Surg Oncol 2015;111:203-7.

105. Yamauchi H, Nakagawa C, Takei H, et al. Prospective study of the effect of the 21-gene assay on adjuvant clinical decision-making in Japanese women with estrogen receptor-positive, node-negative, and node-positive breast cancer. Clinical breast cancer 2014;14:191-7.

106. Fried G, Moskovitz M. Treatment decisions in estrogen receptor-positive early breast cancer patients with intermediate oncotype DX recurrence score results. Springerplus 2014;3:71.

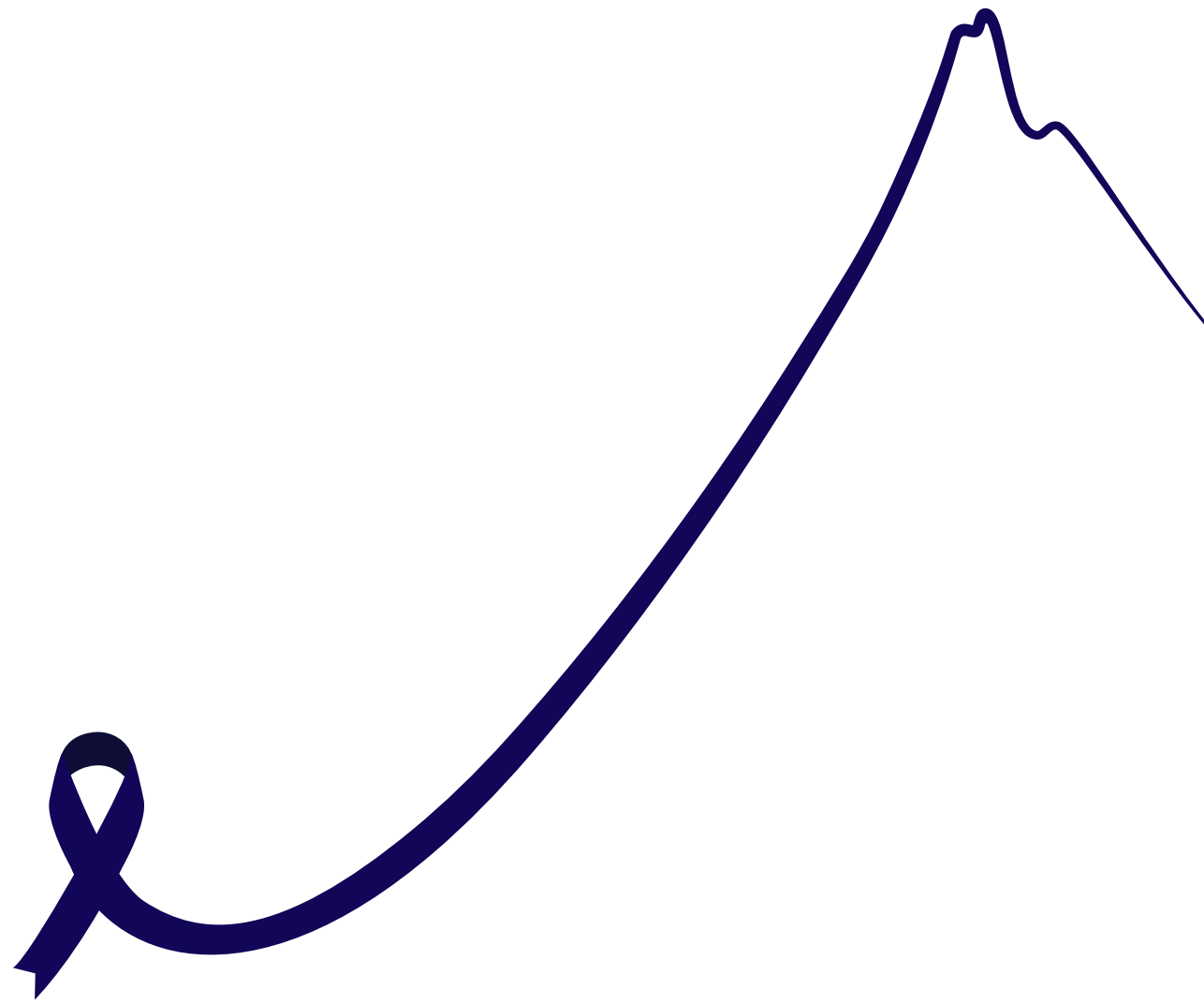
107. Cheung PS, Tong AC, Leung RC, et al. Initial experience with the Oncotype DX assay in decision-making for adjuvant therapy of early oestrogen receptor-positive breast cancer in Hong Kong. Hong Kong Med J 2014;20:401-6.

108. Eiermann W, Rezai M, Kummel S, et al. The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. Annals of Oncology 2013;24:618-24.

109. de Boer RH, Baker C, Speakman D, et al. The impact of a genomic assay (Oncotype DX) on adjuvant treatment recommendations in early breast cancer. Med J Aust 2013;199:205-8.
110. Geffen DB, Abu-Ghanem S, Sion-Vardy N, et al. The impact of the 21-gene recurrence score assay on decision making about adjuvant chemotherapy in early-stage estrogen-receptor-positive breast cancer in an oncology practice with a unified treatment policy. Annals of Oncology 2011;22:2381-6.
111. Martin M, Gonzalez-Rivera M, Morales S, et al. Prospective study of the impact of the Prosigna assay on adjuvant clinical decision-making in unselected patients with estrogen receptor positive, human epidermal growth factor receptor negative, node negative early-stage breast cancer. Current medical research and opinion 2015;31:1129-37.
112. Muller BM, Keil E, Lehmann A, et al. The EndoPredict Gene-Expression Assay in Clinical Practice - Performance and Impact on Clinical Decisions. PLoS One 2013;8:e68252.
113. Su KW, Hall J, Soulos PR, et al. Association of 21-gene recurrence score assay and adjuvant chemotherapy use in the medicare population, 2008-2011. J Geriatr Oncol 2016;7:15-23.
114. Potosky AL, O'Neill SC, Isaacs C, et al. Population-based study of the effect of gene expression profiling on adjuvant chemotherapy use in breast cancer patients under the age of 65 years. Cancer 2015;121:4062-70.
115. Hassett MJ, Silver SM, Hughes ME, et al. Adoption of gene expression profile testing and association with use of chemotherapy among women with breast cancer. Journal of Clinical Oncology 2012;30:2218-26.
116. Stemmer SM, Klang SH, Ben-Baruch N, et al. The impact of the 21-gene Recurrence Score assay on clinical decision-making in node-positive (up to 3 positive nodes) estrogen receptor-positive breast cancer patients. Breast Cancer Res Treat 2013;140:83-92.
117. Ray GT, Mandelblatt J, Habel LA, et al. Breast cancer multigene testing trends and impact on chemotherapy use. Am J Manag Care 2016;22:e153-60.
118. Kuijer A, van Bommel AC, Drukker CA, et al. Using a gene expression signature when controversy exists regarding the indication for adjuvant systemic treatment reduces the proportion of patients receiving adjuvant chemotherapy: a nationwide study. Genet Med 2016;18:720-6.
119. Epstein AJ, Wong YN, Mitra N, et al. Adjuvant Chemotherapy Use and Health Care Costs After Introduction of Genomic Testing in Breast Cancer. Journal of Clinical Oncology 2015;33:4259-+.
120. Katz G, Romano O, Foa C, et al. Economic Impact of Gene Expression Profiling in Patients with Early-Stage Breast Cancer in France. PLoS One 2015;10:e0128880.
121. Reed SD, Dinan MA, Schulman KA, et al. Cost-effectiveness of the 21-gene recurrence score assay in the context of multifactorial decision making to guide chemotherapy for early-stage breast cancer. Genet Med 2013;15:203-11.
122. Vataire AL, Laas E, Aballea S, et al. [Cost-effectiveness of a chemotherapy predictive test]. Bull Cancer 2012;99:907-14.
123. Oestreicher N, Ramsey SD, Linden HM, et al. Gene expression profiling and breast cancer care: what are the potential benefits and policy implications? Genet Med 2005;7:380-9.
124. Bonastre J, Marguet S, Lueza B, et al. Cost effectiveness of molecular profiling for adjuvant decision making in patients with node-negative breast cancer. J Clin Oncol 2014;32:3513-9.
125. Chen E, Tong KB, Malin JL. Cost-effectiveness of 70-gene MammaPrint signature in node-negative breast cancer. Am J Manag Care 2010;16:e333-42.
126. Kondo M, Hoshi SL, Ishiguro H, et al. Economic evaluation of the 70-gene prognosis-signature (MammaPrint(R)) in hormone receptor-positive, lymph node-negative, human epidermal growth factor receptor type 2-negative early stage breast cancer in Japan. Breast Cancer Res Treat 2012;133:759-68.
127. Retel VP, Joore MA, Knauer M, et al. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. Eur J Cancer 2010;46:1382-91.
128. Retel VP, Joore MA, Drukker CA, et al. Prospective cost-effectiveness analysis of genomic profiling in breast cancer. Eur J Cancer 2013;49:3773-9.
129. Bacchi CE, Prisco F, Carvalho FM, et al. Potential economic impact of the 21-gene expression assay on the treatment of breast cancer in Brazil. Rev Assoc Med Bras (1992) 2010;56:186-91.
130. Cosler LE, Lyman GH. Economic Analysis of Gene Expression Profile Data to Guide Adjuvant Treatment in Women with Early-Stage Breast Cancer. Cancer Investigation 2009;27:953-9.
131. Hannouf MB, Xie B, Brackstone M, et al. Cost-effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in women with early-stage estrogen- or progesterone-receptor-positive, axillary lymph-node negative breast cancer. BMC Cancer 2012;12:447.

132. Hornberger J, Cosler LE, Lyman GH. Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. 2005;11:313-24.
133. Hornberger J, Chien R, Krebs K, et al. US Insurance Program's Experience With a Multigene Assay for Early-Stage Breast Cancer. American Journal of Managed Care 2011;17:E194-E202.
134. Jahn B, Rochau U, Kurzthaler C, et al. Cost effectiveness of personalized treatment in women with early breast cancer: the application of OncotypeDX and Adjuvant! Online to guide adjuvant chemotherapy in Austria. Springerplus 2015;4:752.
135. Klang SH, Hammerman A, Liebermann N, et al. Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization. Value Health 2010;13:381-7.
136. Kondo M, Hoshi SL, Ishiguro H, et al. Economic evaluation of 21-gene reverse transcriptase-polymerase chain reaction assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer in Japan. Breast Cancer Research and Treatment 2008;112:175-87.
137. Kondo M, Hoshi SL, Yamanaka T, et al. Economic evaluation of the 21-gene signature (Oncotype DX) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). Breast Cancer Res Treat 2011;127:739-49.
138. Lamond NW, Skedgel C, Rayson D, et al. Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer. Breast Cancer Res Treat 2012;133:1115-23.
139. Health Quality O. Gene Expression Profiling for Guiding Adjuvant Chemotherapy Decisions in Women with Early Breast Cancer: An Evidence-Based and Economic Analysis. Ontario Health Technology Assessment Series 2010;10:1-57.
140. Paulden M, Franek J, Pham B, et al. Cost-Effectiveness of the 21-Gene Assay for Guiding Adjuvant Chemotherapy Decisions in Early Breast Cancer. Value in Health 2013;16:729-39.
141. Smyth L, Watson G, Walsh EM, et al. Economic impact of 21-gene recurrence score testing on early-stage breast cancer in Ireland. Breast Cancer Res Treat 2015;153:573-82.
142. Tsoi DT, Inoue M, Kelly CM, et al. Cost-Effectiveness Analysis of Recurrence Score-Guided Treatment Using a 21-Gene Assay in Early Breast Cancer. Oncologist 2010;15:457-65.
143. Ward S, Scope A, Rafia R, et al. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: a systematic review and cost-effectiveness analysis. Health technology assessment (Winchester, England) 2013;17:1-302.
144. Yamauchi H, Nakagawa C, Yamashige S, et al. Societal cost-effectiveness analysis of the 21-gene assay in estrogen-receptor-positive, lymph-node-negative early-stage breast cancer in Japan. BMC Health Serv Res 2014;14:372.
145. Bargallo-Rocha JE, Lara-Medina F, Perez-Sanchez V, et al. Cost-effectiveness of the 21-gene breast cancer assay in Mexico. Adv Ther 2015;32:239-53.
146. Blohmer JU, Rezai M, Kummel S, et al. Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: a cost-effectiveness evaluation in the German setting. 2013;16:30-40.
147. Hall PS, McCabe C, Stein RC, et al. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. Journal of the National Cancer Institute 2012;104:56-66.
148. Hannouf MB, Xie B, Brackstone M, et al. Cost effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in post-menopausal women with early-stage estrogen or progesterone-receptor-positive, axillary lymph-node positive breast cancer. PharmacoEconomics 2014;32:135-47.
149. Kip M, Monteban H, Steuten L. Long-term cost-effectiveness of Oncotype DX(R) versus current clinical practice from a Dutch cost perspective. J Comp Eff Res 2015;4:433-45.
150. Nerich V, Curtit E, Bazan F, et al. [Economic assessment of the routine use of Oncotype DX(R) assay for early breast cancer in Franche-Comte region]. Bull Cancer 2014;101:681-9.
151. Vanderlaan BF, Broder MS, Chang EY, et al. Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. Am J Manag Care 2011;17:455-64.
152. Blank PR, Filipits M, Dubsky P, et al. Cost-effectiveness analysis of prognostic gene expression signature-based stratification of early breast cancer patients. PharmacoEconomics 2015;33:179-90.
153. Mislick K, Schonfeld W, Bodnar C, et al. Cost-effectiveness analysis of Mammostrat(R) compared with Oncotype DX(R) to inform the treatment of breast cancer. Clinicoecon Outcomes Res 2014;6:37-47.
154. Retel VP, Joore MA, van Harten WH. Head-to-head comparison of the 70-gene signature versus the 21-gene assay: cost-effectiveness and the effect of compliance. Breast Cancer Res Treat 2012;131:627-36.

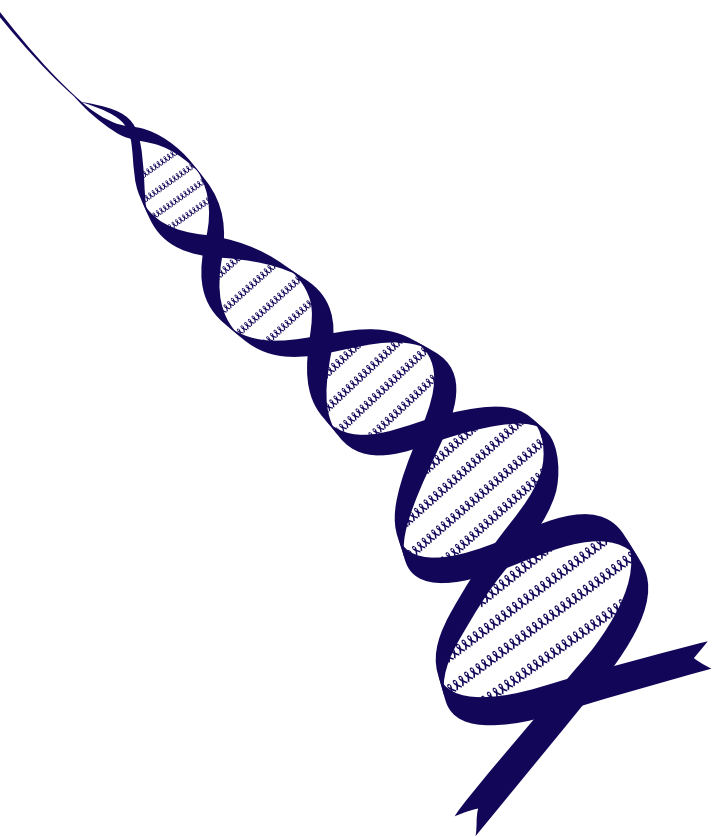
155. Segui MA, Crespo C, Cortes J, et al. Genomic profile of breast cancer: cost-effectiveness analysis from the Spanish National Healthcare System perspective. Expert Rev Pharmacoecon Outcomes Res 2014;14:889-99.
156. Yang M, Rajan S, Issa AM. Cost effectiveness of gene expression profiling for early stage breast cancer: a decision-analytic model. Cancer 2012;118:5163-70.
157. Katz SJ, Morrow M. Addressing overtreatment in breast cancer: The doctors' dilemma. Cancer 2013;119:3584-8.
158. Gnant M, Steger GG. Fighting overtreatment in adjuvant breast cancer therapy. Lancet 2009;374:2029-30.
159. Petkov VI, Miller DP, Howlader N, et al. Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. npj Breast Cancer 2016;2:16017.
160. Hornberger J, Alvarado MD, Rebecca C, et al. Clinical validity/utility, change in practice patterns, and economic implications of risk stratifiers to predict outcomes for early-stage breast cancer: a systematic review. Journal of the National Cancer Institute 2012;104:1068-79.
161. Marrone M, Stewart A, Dotson WD. Clinical utility of gene-expression profiling in women with early breast cancer: an overview of systematic reviews. Genetics in medicine : official journal of the American College of Medical Genetics 2015;17:519-32.
162. Esserman LJ, Yau C, Thompson CK, et al. Use of Molecular Tools to Identify Patients With Indolent Breast Cancers With Ultralow Risk Over 2 Decades. Jama Oncol 2017.
163. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet 2015;386:1341-52.



Chapter 10

Letter to the editor: 70-Gene Signature in Early-Stage Breast Cancer

E.J. Blok
C.J.H. van de Velde
V.T.H.B.M. Smit



New England Journal of Medicine
2016; 375:2199-2201

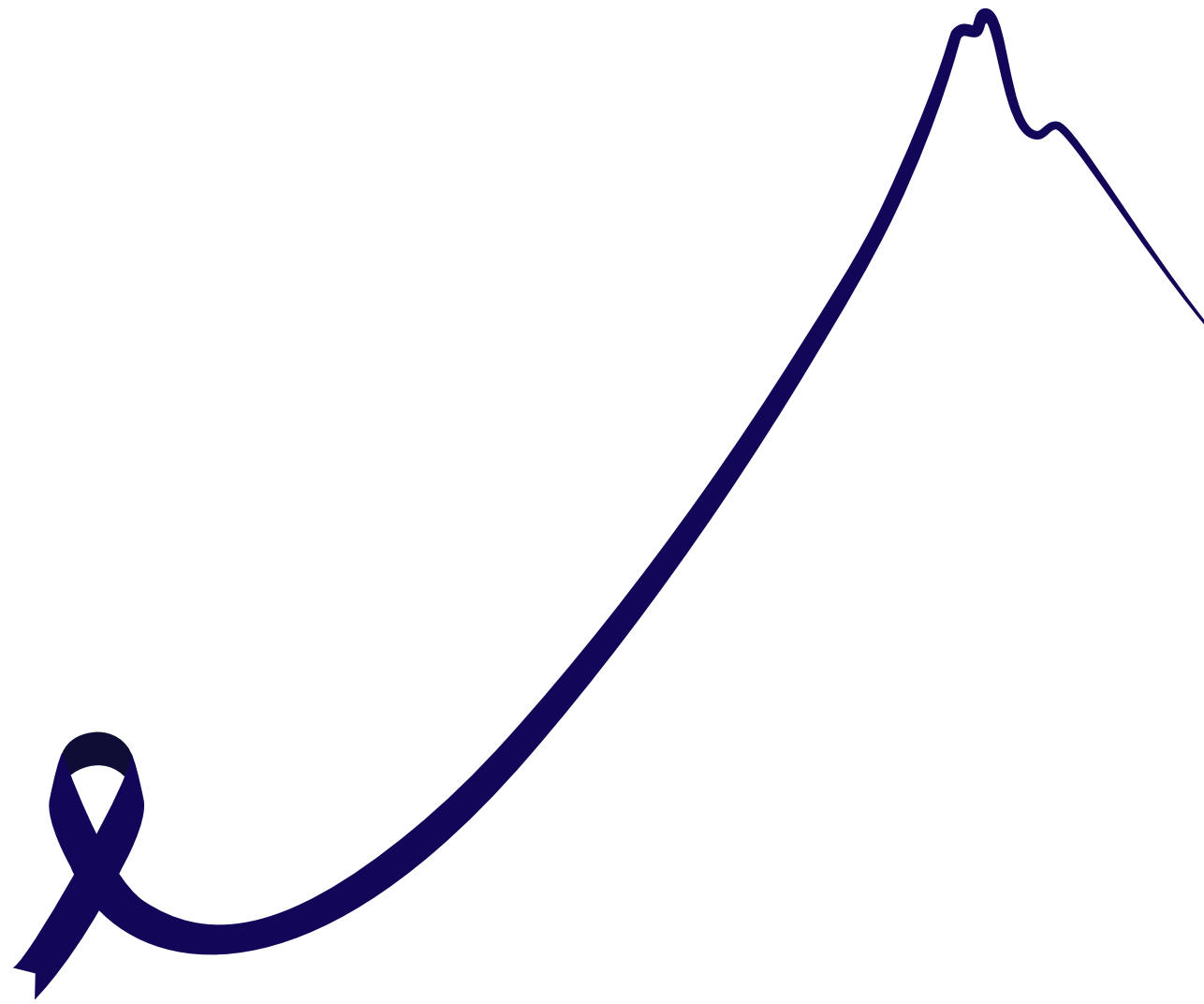
The results of the study by Cardoso et al. (Aug. 25 issue)[1] suggest that chemotherapy can be safely withheld from patients who are clinically at high risk for recurrence but have a low-risk 70-gene signature (MammaPrint). However, the subgroup analysis does not show whether this finding was also true for patients with grade 3 tumors (found in 29% of the patients), who usually have an increased benefit from chemotherapy. Besides this factor, an underexposed finding of this study is that MammaPrint was not useful in at least 60% of the patients, particularly those at low clinical risk and those at high clinical risk with triple-negative tumors.

There were major differences between the characteristics of the patients at high clinical risk but low genomic risk and the characteristics of those at high risk in both categories. Among patients at high clinical risk but low genomic risk, 90% of the tumors were luminal and negative for human epidermal growth factor receptor 2 (HER2), and 71% of the tumors were grade 1 or 2. In contrast, among the patients at high clinical and genomic risk, only 50% of the tumors were luminal and HER2-negative, and 76% were grade 3. We calculated that among the patients at high clinical risk, 82% of luminal grade 1 or 2 tumors would be classified as genomic low risk.

Genomic assays are expensive and should be used efficiently. It may be possible to perform a decision-tree analysis on the basis of chi-square automatic interaction detection (CHAID)[2] using primary intrinsic tumor characteristics (e.g., the presence or absence of estrogen receptor and HER2, along with Ki-67 status and tumor grade) as predictors for the MammaPrint outcome. On the basis of the outcome of such a study, the use of MammaPrint could be restricted to patients for whom the clinicopathological risk assessment is insufficient.

References

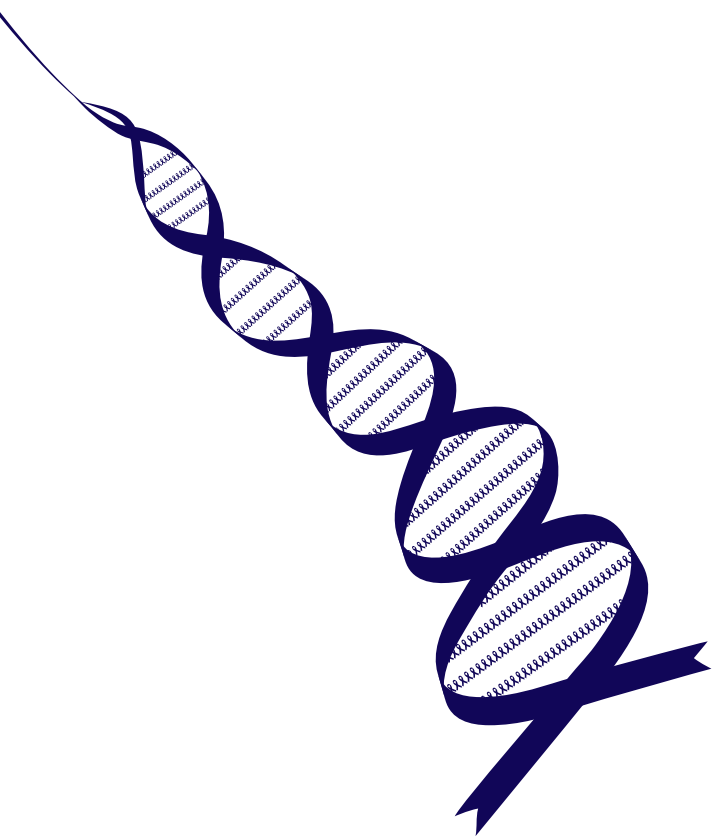
1. Cardoso F, Van't Veer LJ, Bogaerts J et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med 2016;375(8):717-729.
2. Kass GV. An Exploratory Technique for Investigating Large Quantities of Categorical Data. Journal of the Royal Statistical Society Series C (Applied Statistics) 1980;29(2):119-127.



Chapter 11

General discussion

E.J. Blok



In the past decades, adjuvant endocrine therapy has increased substantially in popularity and has become one of the mainstays in the treatment of patients with breast cancer. Current international guidelines state that all patients with >1% of tumour cells expressing ER are eligible to receive adjuvant endocrine therapy.[1] Although the Dutch guidelines are a bit less stringent (>10% ER expression, and <2 cm and grade 1, or <1cm and grade 1 or 2 tumours are excluded), still the majority of ER-positive patients receives adjuvant endocrine therapy.[2-4] The treatment durations have escalated over the years. Starting initially with just a few months of adjuvant tamoxifen, trials have been performed studying up to 15 years of adjuvant endocrine therapy.[5]

Both the relative low threshold for ER-positivity and the increasing treatment durations contribute to the risk of overtreatment. In a meta-analysis performed by the Early Breast Cancer Trialists Collaborative Group (EBCTCG), 5 years of tamoxifen was compared to no adjuvant therapy. At 15 years after diagnosis, there was an absolute benefit of 12% in recurrences (45% vs 33%), and an absolute benefit of 9% (35% vs 26%) on breast cancer mortality.[6] However, this also suggests that for 88% of the patients there was no benefit of tamoxifen (55% because they wouldn't develop a metastasis anyhow, and 33% because they still develop a metastasis despite their therapy). Identification of these patients is crucial, either to de-escalate therapy for the ones that would not develop a metastasis even without therapy, or to escalate therapy for the patients that would develop a (late) metastasis despite their therapy. In this thesis, we have studied and discussed multiple aspects of tailoring this adjuvant endocrine therapy.

Extended adjuvant endocrine therapy

Extended therapy beyond 5 years is one of the strategies to escalate therapy, in order to prevent late relapses of HR-positive breast cancer. Especially after 5 years of tamoxifen, it is often considered to be standard of care, which is reflected in most international guidelines and is summarized in chapter 2. The leading study in this field is the MA.17 trial, which was published in 2003 by the group of Paul Goss in the New England Journal of Medicine.[7] In this study, over 5000 women who earlier received 5 years of tamoxifen, were randomized between 5 years of extended letrozole, or 5 years of placebo. At interim analysis after 2.4 years, the disease-free survival in the treated group was 93%, versus 87% in the placebo group (HR 0.57, 95% CI 0.43-0.75, p<0.001). Although impressive at first glance, the absolute differences in terms of

distant recurrences are less impressive. In the letrozole group 47 (out of 2593) patients had a distant recurrence (1.8%), in the placebo group 76 (out of 2594) had a distant recurrence (2.9%). Based on these absolute numbers, it can be questioned whether this 1.1% of absolute difference in distant metastasis justifies 5 years of additional therapy. This doubt is strengthened by the 5 year follow-up publication, which showed no significant difference in distant metastasis-free survival (HR 0.80, 95% CI 0.62-1.03, p=0.08) or overall survival (HR 0.98, 95% CI 0.78-1.22, p=0.85).[8] That DFS was still significantly improved in both analyses, is explained by the fact that death due to other causes was not included in the definition of DFS. Furthermore, prevention of local relapse and secondary breast cancer in the contralateral breast (which is not regarded as a treatment aim for systemic adjuvant therapy) also strongly influenced the differences in DFS.

In a recent meta-analysis, it was confirmed that in general the added effect of extended endocrine therapy is limited, especially when overall survival is used as outcome measure.[9] For recurrences, the same meta-analysis shows that the effect is isolated to patients with positive lymph nodes.[9] In chapter 3 of this thesis, we describe the results of the phase III IDEAL trial, in which postmenopausal patients with early HR+ breast cancer were randomized between either 2.5 or 5 years of letrozole, after finishing 5 years of regular adjuvant endocrine therapy. In this chapter, we conclude that longer (5 versus 2.5 years) extended therapy has little value for the full population (chapter 3). However, other groups studying extended endocrine therapy, suggested that for patients with node-positive disease, there might be a benefit of longer AI therapy after using sequential therapy of tamoxifen followed by an AI for 5 years.[10] In chapter 4, we describe a subgroup analysis in the IDEAL trial, in this particular subgroup (node-positive disease, pre-treated with tamoxifen followed by an AI). In this chapter, we have shown that a longer use of letrozole in this particular subgroup might be beneficial. Still, despite the significant value in node-positive disease, the absolute benefits of extended therapy remain small. Therefore, shared decision-making between patients and physicians plays a major role, balancing the (small) benefits and side effects.

Another reason why shared decision-making is particularly important for extended endocrine therapy, is the compliance to therapy. In the primary analysis of the IDEAL trial (chapter 3), we have shown that 25% of patients in the 2.5 years, and 45% in the 5 years group are unable to finish therapy, in majority explained by adverse events.

In chapter 5, we further investigated this phenomenon, by evaluating the factors associated to participating in the IDEAL, the factors associated to early treatment discontinuation, and the effect of early treatment discontinuation on survival outcome. We showed that factors associated to participation are high risk factors like a younger age and nodal status, whereas the factors associated to early discontinuation are more patient-centred factors like the type of earlier endocrine therapy, the amount of time between treatments, and the occurrence of side effects. Remarkably, we have shown that patients who decide to cease therapy after an adverse event, have an equal survival outcome compared to those who continue with therapy after an adverse event. This emphasizes the need for shared decision-based, personalized treatment regimes.

One of the outcomes that shows a consistency under extended endocrine therapy, is the lower occurrence of contralateral breast cancer, which was also shown in chapter 3. This preventive effect of endocrine therapy on the occurrence of new primary breast tumours is well studied, and has increasing popularity.[11] However, the differences in absolute and relative risk reductions play a major role in this discussion. One of the most well-known studies in the field of primary breast cancer prevention is the International Breast Intervention Study II (IBIS-II). They randomized 3864 patients between anastrozole or placebo. After 5 years, there was a 50% reduction in the incidence of breast cancers (32 vs 64 respectively). This absolute reduction of 32 cases represents an absolute decrease of 1.7%, which already sounds much less impressive. Combined with the fact that only ER-positive breast cancer is prevented, which in general has a more favourable prognosis (approximately 85-90% survival at 5 years), the effect on overall survival is almost non-existent. Therefore, we feel that the preventive effect of extended endocrine therapy should not be used as an argument for the use of (extended) endocrine therapy.

Biomarker-based personalized endocrine therapy

One approach to improve the effect of (extended) endocrine therapy is to identify the patients that will benefit most from it, using biomarkers. Or, vice versa, use biomarkers to identify the patients that will not benefit from it, so that other types of therapy can be considered. Roughly, there are two approaches in the development of biomarkers, which we will call biology-based and risk-based biomarkers.

The first approach, *biology-based biomarkers*, are biomarkers that are designed based on the biological mechanism of an intervention. A current example which is already widely used, is the tumor expression of hormone receptors (HRs). When these receptors are not expressed, endocrine therapy is not expected to cause any therapeutical effect. However, as shown in the previous section, the expression of hormone receptors is not a guarantee for treatment success. A possible mechanism to improve the use of information on the tumour expression of hormone receptors as predictive biomarkers for endocrine therapy, is described in chapter 6. By determining the activity of the ER-pathway, you could distinguish for which patients the estrogen receptor is not only expressed, but indeed active and therefore a suitable target for therapy. In chapter X, we adapted this procedure for evaluation in the TEAM IIA trial, in which patients were treated with neo-adjuvant endocrine therapy, we showed that non-response and progressive disease during therapy were associated to a lower baseline ER-pathway activity. Furthermore, in a public dataset, the decrease in ER-pathway activity was associated to therapy response. Therefore, this technique might be a way to monitor the efficacy of endocrine therapy, since the receptor pathway activity is expected to diminish upon successful treatment. Currently this is only applicable to the neo-adjuvant and metastatic setting, since these are the only settings with a tumour in situ for monitoring. However, with increasing utility of circulating tumour cells, this technique might become feasible for adjuvant therapies as well. Furthermore, future analyses are planned to assess whether a lack of decrease in ER-pathway activity might be explained by baseline mutations in ESR1, the gene coding for ER. If this is the case, patients with such a mutation might be spared from endocrine therapy since they will have no clinical response, and they should be treated with other types of adjuvant therapy (i.e. chemotherapy).

A second biology-based biomarker approach highlighted in this thesis, is the tumour-immune environment, specifically the *tumour-infiltrating lymphocytes* (TILs). TILs, and specifically CD8-positive TILs are effector cells of the adapted immune system, capable of targeting tumour cells which they recognize as being 'foreign' due to expression of tumour neo-epitopes. However, since TILs are depending on these neo-epitopes for their activation, tumours with a lower mutational load are usually considered to be less responsive against TIL-infiltration.

It has been shown that ER-positive tumours have a lower mutational load compared to ER-negative tumours.[12]o Therefore, it is no surprise that TILs have no prognostic

value in ER-positive disease, in contrast to ER-negative disease in which high numbers of TILs predict for a better survival.[13, 14] The lack of prognostic value of TILs in ER-positive disease was confirmed in this thesis in multiple cohorts (chapter 7 and 8). In contrast, TILs have prognostic capacities in ER-negative disease, in particular in triple-negative breast cancer (TNBC). [15-19] In chapter 8 we explored the role of FAS, a key mediator in cytotoxic T-cell based immunity, in the distinction between ER-positive and ER-negative disease. We showed that CD8-positive TILs only had prognostic value in the presence of FAS expression , and that FAS was expressed twice as frequent in ER-negative disease compared to ER-positive disease.

In chapter 7, we evaluated the predictive capacity of CD8-positive TILs in the Dutch population of the Intergroup Exemestane trial (IES), which randomized patients between tamoxifen or exemestane after 2-3 years of tamoxifen. In this analysis, we have shown a strong predictive value of TILs in ER-positive disease, with regard to a differential treatment response to either tamoxifen or an AI. Patients with low numbers of TILs had a more favourable prognosis when treated with an AI compared to tamoxifen, whereas patients with a high number of TILs had a similar prognosis on both treatments.

We have two different hypotheses for this observation. A first explanation might be a direct influence of endocrine therapy on lymphocytes in general, and TILs in particular. It is known that ER is expressed in lymphocytes, and the response in these cells to estrogen depletion (with an AI) might be different from the response to receptor modulation by tamoxifen.[20] This could theoretically lead to altered functionality of the TILs, and thereby a difference in clinical prognosis. The second theory to explain the findings of TILs as predictive markers for endocrine therapy, could be that the number of TILs are a proxy marker for the mutational load. Tumours with higher numbers of TILs have a higher mutational load[21], more resembling ER-negative tumours and less dependent on ER-signalling. In that case, the type of endocrine therapy would make little difference. In contrast, tumours with lower levels of TILs may have a lower mutational load, thereby being more dependent on ER-signalling. This strong ER dependency might magnify the differences between AIs and tamoxifen with regard survival benefits. Future studies will need to show which of these two theories explains our results best, and validation is required before this marker can be used in a clinical setting.

Risk-based biomarkers

Risk-based biomarkers are capable of discriminating between patients with a high or low risk of tumour recurrence. Even when the relative treatment benefit (hazard ratio) is equal in the low-risk and high-risk subgroup, the treatment will have an higher impact when the a priori chance of recurrence is higher. For example, when a group of patients with a 50% chance and a group with a 10% chance of recurrence a being treated with a therapy that has a hazard ratio of 0.5, the first group will have an absolute risk reduction of 25% (1 in 4 patients has a benefit), whereas the second group has an absolute risk reduction of only 5% (1 in 20 patients has a benefit). Therefore, selection of either high-risk or low-risk patients might help in selective escalating and de-escalating of endocrine therapy.

One of the most popular new strategies to identify patients with a lower or higher risk of recurrences, is the use of gene expression profiles (GEPs). These assays determine the risk of recurrence, based on the expression of selected genes in the tumour. These assays are thoroughly discussed in chapter 9 and 10. In chapter 9, we performed an elaborate systematic review, to assess the assay development, clinical validation, clinical utility, and economic value of the four most frequently used GEPs in Europe. In this review, we conclude that in particular OncotypeDX and Mammaprint are both well studied, having level IA evidence available from large randomized trials. In chapter 10, we comment on MINDACT, one of these large trials assessing the clinical functioning of the Mammaprint test together with traditional clinicopathological guidelines. In this letter, we emphasize the need for subgroup analysis and the selection of patients for which testing is the most beneficial, and ask for careful interpretation of the trial results.

The use of GEPs to select patients for endocrine therapy is still limited. Only recently, a relative small study using GEP in an old trial (1976-1990) randomizing between 2 years of tamoxifen and no endocrine therapy, showed that a group of ER-positive patients with ultralow risk had an excellent prognosis, even without endocrine therapy.[22] Upon validation, this or similar other assays could be used to identify the patients for who endocrine therapy can be safely withheld. On the other end, these assays could perhaps be used to identify the patients with a higher risk for tumour recurrence, who might benefit from extended endocrine therapy. This use will be the topic of further investigations, both in the IDEAL trial and in other studies.

Future perspectives

The trend of personalized and precision medicine in oncology is unstoppable, and will change the field completely. The field of breast cancer was one of the first to adopt personalized targeted medicine with endocrine and HER2-targeted therapy. This thesis has shown that for endocrine therapy a further personalization is likely and, upon validation of our findings, will lead to a more optimal treatment for every individual patient. However, there are some challenges which will need to be addressed before personalized endocrine therapy will become standard of care.

The first challenge will be to validate the initial results in a way, that can reliably be applied in the clinical setting. In the current situation, in which endocrine therapy regimes only become longer, especially de-escalation will be challenging. Prospective-retrospective studies, in which an earlier randomized trial is used to assess the predictive capacity of a new biomarker, is a popular method to validate a biomarker for treatment decisions. However, trials with ER-positive breast cancer, without any endocrine therapy in one arm (which would be needed to show the safety of biomarker-based de-escalation) are rare and usually old.[6] It can be questioned whether these cohorts are still representative enough for current practice. The validation of biomarker-based differentiation between tamoxifen and AIs might be easier, since these trials (like BIG 1-98 and ATAC) are more recent and are suited for validation.

A second challenge is the remaining risk of tumour recurrence beyond 5 years of standard endocrine therapy.[23] This thesis has shown that extending adjuvant therapy is not the solution for this problem, and there remains a continuous risk despite the extended therapy. A possible explanation might be that endocrine therapy is considered as cytostatic treatment, slowing down or stopping tumour growth without actually inducing cell death. Therefore, one extra step has to be taken in the field of endocrine therapy in order to use it as cytotoxic therapy. Whether this step will be taken using immunotherapy or ER-targeted cytotoxic therapy is not clear yet, and will take many years to develop.

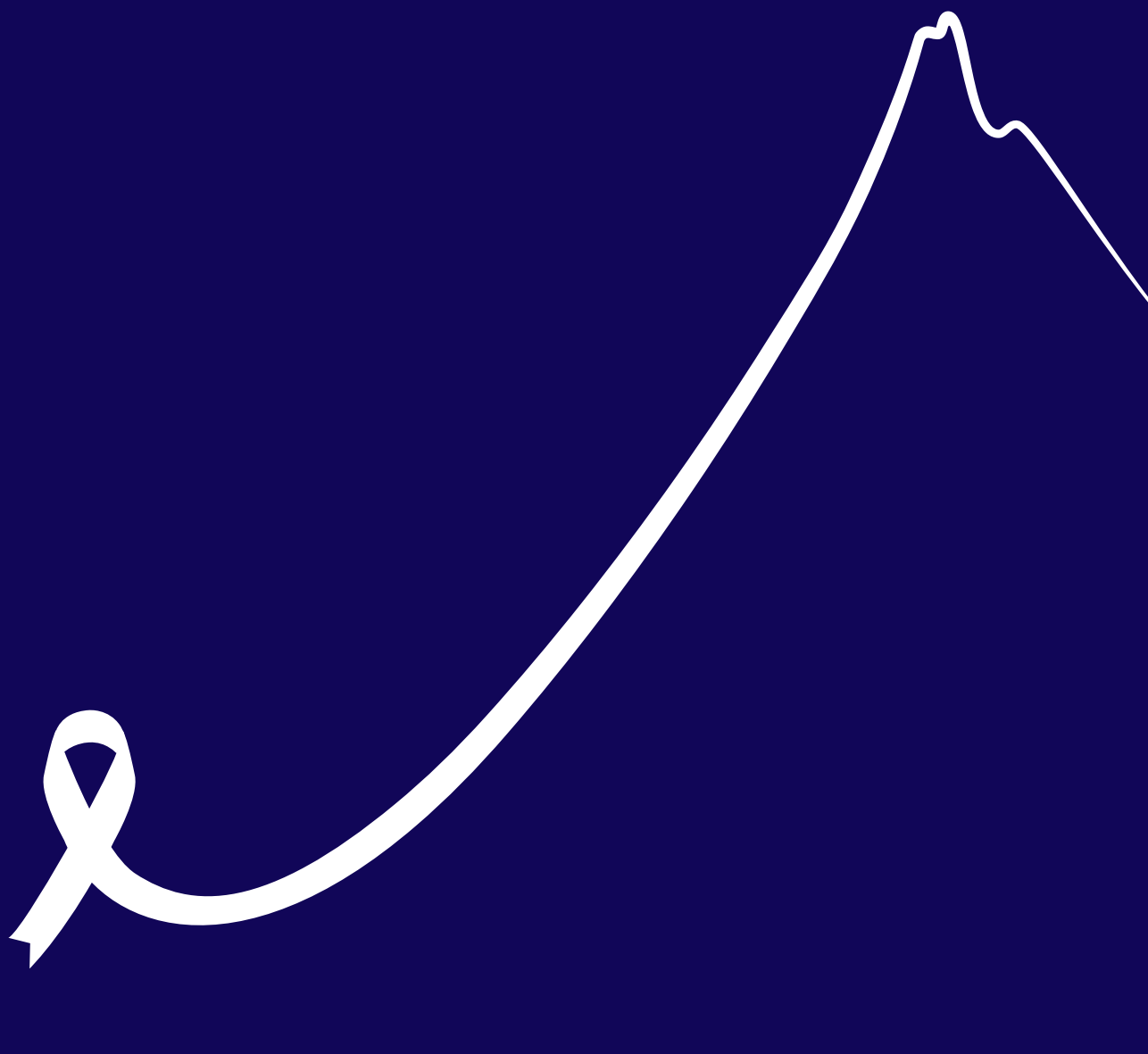
Another challenge, which applies for personalized endocrine therapy but also for personalized medicine in general, is the validation of personalization as the most optimal treatment strategy. In classic evidence-based medicine, trials with thousands of patients are conducted, in which one arm is using the new therapy, and the other

arm is using standard therapy. However, the ultimate goal of personalized medicine, is that every treatment strategy is unique for every patient, and per definition therefore cannot be validated using 'regular' clinical trials. A work-around for this problem might be the development of trials that validate a treatment concept, instead of an individual therapy. In that case, you might randomize between 'therapy according to biomarker-protocol' and 'therapy according to standard protocol'. However, financing this kind of trials for personalized endocrine therapy conserving patient numbers would become complicated, since all (adjuvant) endocrine therapy agents are off-patent, and benefits for industry would be low. For the development of new personalized agents (e.g. combination inhibitors, tailored to the tumour molecular make-up), other problems arise when this new concept of protocol-based treatment validation would be used. In order to be accepted to the US and EU markets, individual agents now have to be registered with evidence from a registration trial, in which the new drug is showing superiority over standard therapy. However, when these agents are tailored to individual patients, these trials are impossible to conduct. The unregistered use of agents in protocol-based trials as described above, would therefore require a paradigm shift in evidence-based medicine and pharmaceutical regulations.

When these challenges are met, the future of endocrine therapy will become a personalized treatment strategy combining targeted cytostatic and cytotoxic approaches. Decisions whether endocrine therapy should be started will be made using clinical and genomic risk evaluations, whereas decisions which type of endocrine therapy will be most effective will be made using biology-based biomarkers like tumour lymphocyte infiltration and ER pathway activity. Combining both approaches will lead to more effective endocrine therapy strategy for every individual patient. Only when we are able to select the most optimal endocrine therapy, we will be able to determine the optimal duration for each approach. Until then, 5 years of adjuvant therapy is sufficient for the majority of patients, and extended endocrine therapy should only be considered for a small subgroup of high-risk, tamoxifen treated patients.

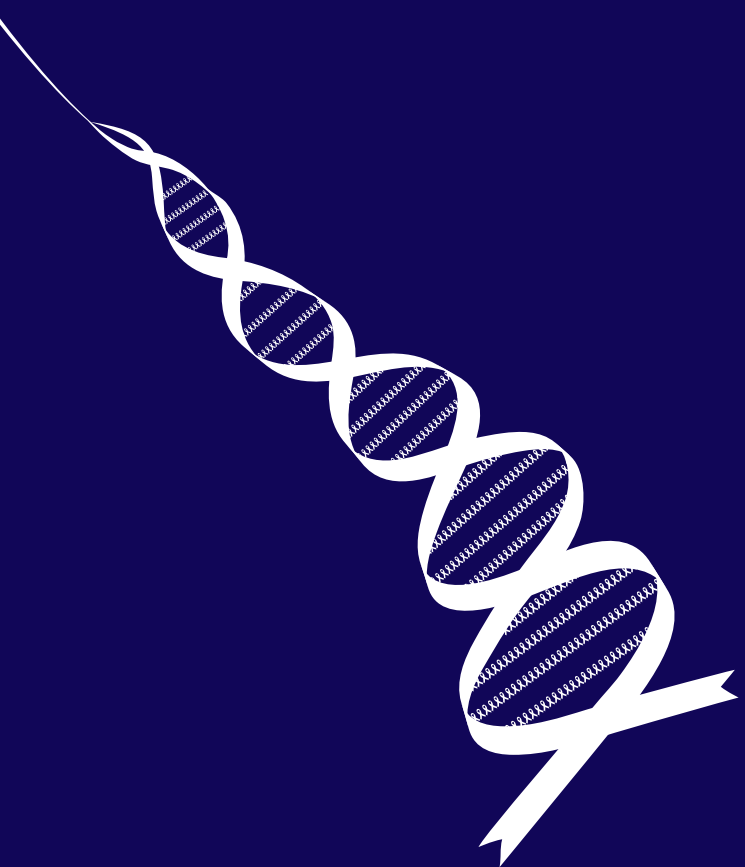
References

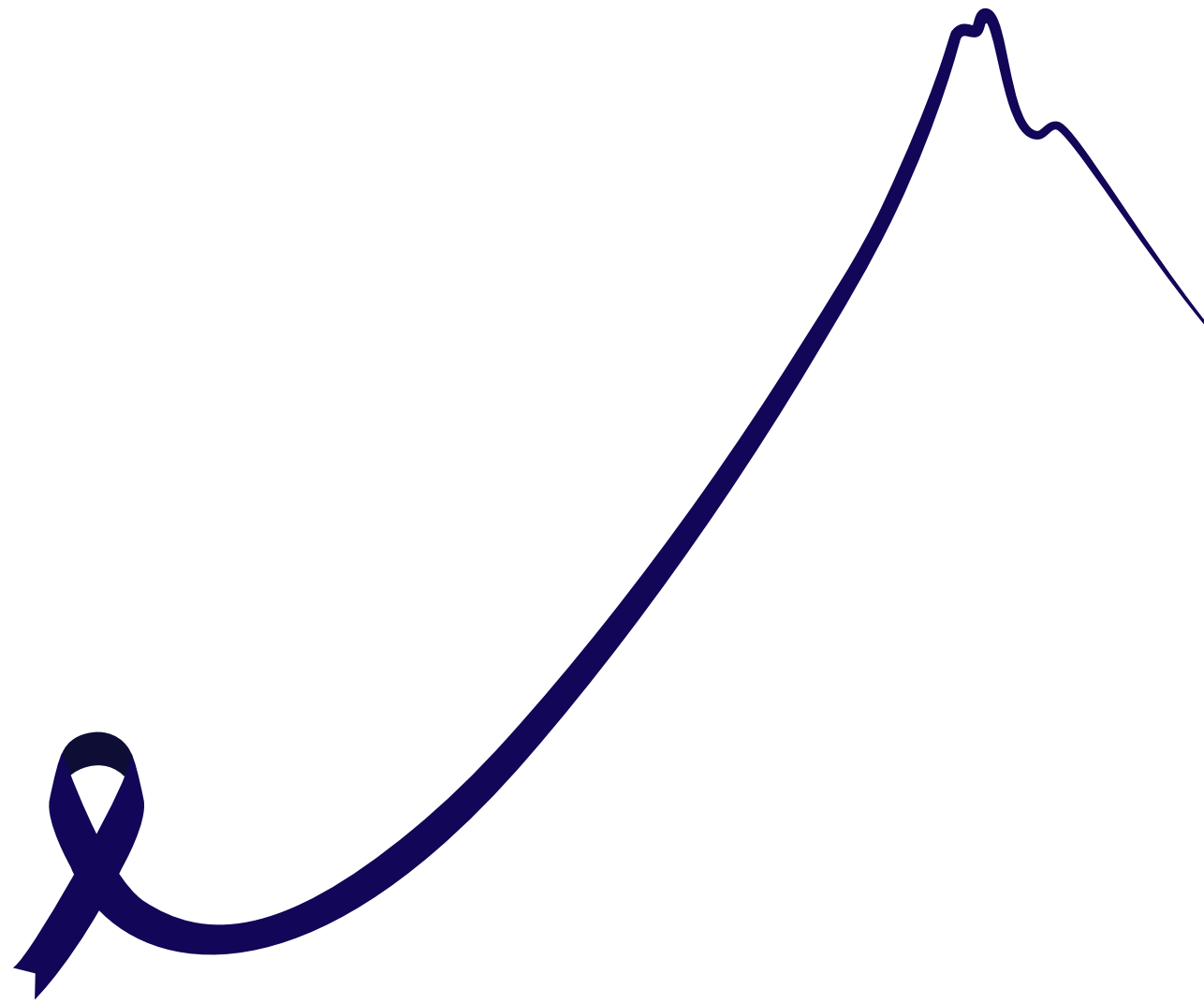
1. Hammond ME, Hayes DF, Dowsett M et al. American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28:2784-2795.
2. NABON. Richtlijn Mammacarcinoom. 2012. 23-12-2015. Ref Type: Internet Communication
3. Kiderlen M, Walsh PM, Bastiaannet E et al. Treatment Strategies and Survival of Older Breast Cancer Patients: An International Comparison between the Netherlands and Ireland. PLoS ONE 2015;10(2):e0118074.
4. IKNL. Dutch Cancer Registry. 14-12-2017. https://www.iknl.nl/oncologische-zorg/tumorteams/borstkanker#incidentie
5. Goss PE, Ingle JN, Pritchard KI et al. Extending Aromatase-Inhibitor Adjuvant Therapy to 10 Years. N Engl J Med 2016;375(3):209-219.
6. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-1717.
7. Goss PE, Ingle JN, Martino S et al. A Randomized Trial of Letrozole in Postmenopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. N Engl J Med 2003;349(19):1793-1802.
8. Ingle JN, Tu D, Pater JL et al. Intent-to-treat analysis of the placebo-controlled trial of letrozole for extended adjuvant therapy in early breast cancer: NCIC CTG MA.17. Ann Oncol 2008;19:877-882.
9. Ibrahim EM, Al-Hajeili MR, Bayer AM et al. Extended adjuvant endocrine therapy in early breast cancer: a meta-analysis of published randomized trials. Medical Oncology 2017;34(7):131.
10. Tjan-Heijnen VCG, van Hellemond IEG, Peer PGM et al. Extended adjuvant aromatase inhibition after sequential endocrine therapy (DATA): a randomised, phase 3 trial. The Lancet Oncology 2017;18(11):1502-1511.
11. Thorat MA, Cuzick J. Preventing invasive breast cancer using endocrine therapy. The Breast 2017;34(Supplement 1):S47-S54.
12. Haricharan S, Bainbridge MN, Scheet P et al. Somatic mutation load of estrogen receptor-positive breast tumors predicts overall survival: an analysis of genome sequence data. Breast Cancer Res Treat 2014;146(1):211-220.
13. Mahmoud SMA, Paish EC, Powe DG et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. JCO 2011;29(15):1949-1955.
14. Ali HR, Provenzano E, Dawson SJ et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. Annals of Oncology 2014;25(8):1536-1543.
15. Denkert C, Loibl S, Noske A et al. Tumor-Associated Lymphocytes As an Independent Predictor of Response to Neoadjuvant Chemotherapy in Breast Cancer. JCO 2010;28(1):105-113.
16. Mahmoud SMA, Paish EC, Powe DG et al. Tumor-Infiltrating CD8+ Lymphocytes Predict Clinical Outcome in Breast Cancer. JCO 2011;29(15):1949-1955.
17. Ono M, Tsuda H, Shimizu C et al. Tumor-infiltrating lymphocytes are correlated with response to neoadjuvant chemotherapy in triple-negative breast cancer. Breast Cancer Res Treat 2012;132(3):793-805.
18. Loi S, Sirtaine N, Piette F et al. Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. JCO 2013;31(7):860-867.
19. Ibrahim E, Al-Foheidi M, Al-Mansour M et al. The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. Breast Cancer Res Treat 2014;148(3):467-476.
20. Scariano JK, Emery-Cohen AJ, Pickett GG et al. Estrogen receptors alpha (ESR1) and beta (ESR2) are expressed in circulating human lymphocytes. Journal of Receptors and Signal Transduction 2008;(1079-9893 (Print)).
21. McGranahan N, Furness AJ, Rosenthal R et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351(6280):1463-1469.
22. Esserman LJ, Yau C, Thompson CK. Use of molecular tools to identify patients with indolent breast cancers with ultralow risk over 2 decades. JAMA Oncology 2017.
23. Pan H, Gray R, Braybrooke J et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. N Engl J Med 2017;377(19):1836-1846.



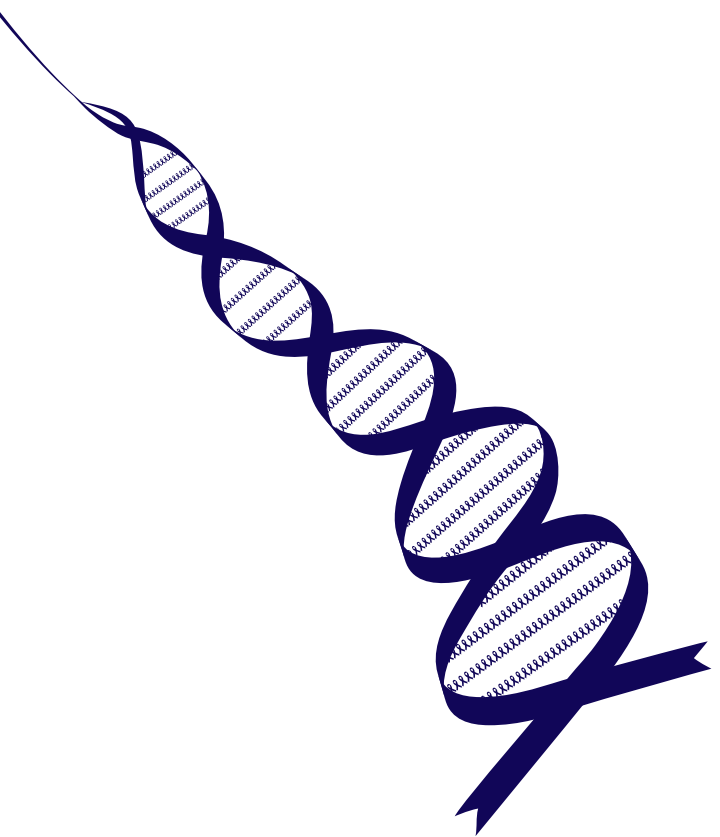
Part 3

Appendices





Nederlandse samenvatting
List of publications
List of co-authors
Curriculum Vitae
Dankwoord



Nederlandse samenvatting

Borstkanker is na longkanker de meest voorkomende soort kanker, en met 17.000 gevallen per jaar de meest voorkomende soort kanker bij vrouwen in Nederland. De behandeling van borstkanker is gebaseerd op drie pijlers: chirurgie om de primaire tumor te verwijderen, radiotherapie om locoregionale uitbreiding te voorkomen of te behandelen, en systemische therapie om uitzaaiingen op afstand te voorkomen of te behandelen.

De mogelijkheden voor systemische therapie worden steeds uitgebreider. Over het algemeen wordt er onderscheid gemaakt tussen chemotherapie, targeted therapie, bijvoorbeeld gericht tegen de HER2-receptor, en hormoontherapie. De rationale achter hormoontherapie bestaat uit het feit dat in ongeveer 80% van alle borstkankers, de oestrogeenreceptor (ER) of progesteronreceptor (PR) tot expressie komt. De tumor 'gebruikt' deze expressie om zichzelf groeisignalen te geven. Het doel van hormoontherapie is deze activatie te voorkomen.

De eerste vermelding van hormoontherapie ontstond in 1896, toen dr. Beatson operatief de eierstokken verwijderde bij een vrouw met uitgezaaid borstkanker. Vervolgens bleken de uitzaaiingen kleiner te worden. Vanaf dat moment is duidelijk dat er een hormonale invloed is op borstkanker, en werd een ovariëctomie standaardbehandeling bij borstkanker. Pas veel later, vanaf de jaren 70, kwamen er ook medicinale oplossingen voor hormoontherapie. In eerste instantie werd tamoxifen ontwikkeld, wat de oestrogeenreceptor (deels) blokkeert. In het begin werd dit alleen gebruikt voor de behandeling van uitzaaiingen, maar later werd ook via diverse klinische trials ontdekt dat een preventieve (adjuvante) behandeling met tamoxifen ook uitzaaiingen kan voorkomen. Lange tijd is 5 jaar therapie hiervoor de standaard behandelduur geweest.

In de tussentijd is er ook een tweede klasse van hormoontherapie ontwikkeld, de zogeheten aromataseremmers (AI). Deze remmen het enzym aromatase, wat er normaal voor zorgt dat mannelijke hormonen worden omgezet in de vrouwelijke oestrogenen. Door dit enzym te blokkeren wordt dus de aanmaak van oestrogeen voorkomen. Ook van deze klasse is aangetoond dat deze metastasen kan voorkomen, en met 5 jaar behandeling zelfs nog wat beter dan met 5 jaar tamoxifen.[1] Ook de behandelswitch van tamoxifen naar een AI halverwege de 5 jaar is beter dan 5 jaar tamoxifen.[1]

Ondanks het succes van de hormoontherapie, is de behandeling nog steeds niet optimaal. Ongeveer 20% van de patiënten heeft baat bij hormoontherapie, wat een groot aantal is. Echter, dit betekent ook dat ongeveer 80% voor niets wordt behandeld: 60% omdat ze toch geen uitzaaiingen zouden krijgen ook als ze niet behandeld zouden worden, en nog eens 20% omdat er ondanks de behandeling alsnog uitzaaiingen ontstaan.[2] Er zijn twee belangrijke knelpunten: Wie hebben er precies wel of geen baat van de hormoontherapie? En wat is de optimale behandelduur?

Wat betreft de optimale patiëntselectie, vind er op dit moment een selectie plaats op basis van de aankleuring van ER door de patholoog. Internationaal is de richtlijn dat indien de ER wordt gezien in meer dan 1% van de tumorcellen, de tumor wordt beschouwd als hormoongevoelig, en de patiënt in principe in aanmerking komt voor hormoontherapie. In Nederland zijn we al iets strenger, en houden we 10% aan en moet er ook sprake zijn van enkele andere ongunstige tumorkarakteristieken (zoals tumorgrootte, lymfeklierstatus en tumorgraad). Echter, er zijn veel patiënten die sowieso geen uitzaaiing zouden hebben gekregen, en er zijn ook nog steeds patiënten die ondanks de therapie alsnog een uitzaaiing krijgen. Van deze patiënten zou je kunnen zeggen dat ze onnodig zijn behandeld met hormoontherapie. Daarom is het belangrijk om van tevoren te kunnen voorspellen welke patiënten dit betreft, zodat je ze gerichter kunt behandelen.

Wat betreft de optimale behandelduur, is 5 jaar adjuvante therapie lange tijd de standaard geweest. Echter, we weten ook dat meer dan de helft van alle recidieven, ná de eerste 5 jaar van follow-up valt. Daarom lijkt het niet onlogisch om ook de adjuvante therapie langer te maken, om die late recidieven ook te voorkomen. Na 5 jaar tamoxifen, is al in meerdere studies aangetoond dat een behandeling tot 10 jaar (of met tamoxifen, of met een AI) beter is dan 5 jaar[3-5], alhoewel dit effect beperkt lijkt tot patiënten met een relatief hoog risico (met lymfekliermetastasen).[6] Voor de behandeling van 5 jaar met een aromataseremmer, is het op dit moment nog niet duidelijk of langere therapie zinvol is en zo ja, voor welke subgroep patiënten.

In dit proefschrift wordt ingegaan op deze twee belangrijke aspecten om de behandeling met hormoontherapie te optimaliseren. In de eerste hoofdstukken ga ik in op de verlengde hormoontherapie, terwijl ik in de latere hoofdstukken me richt op het gebruik van biomarkers om de patiënten te selecteren die het meeste baat hebben van adjuvante hormoontherapie.

Verlengde hormoontherapie

Hoofdstuk 2 behandelt al het huidige bewijs wat er voor aanvang van dit proefschrift bestond voor verlengde hormoontherapie. Hierin wordt er getoond dat tot nu toe er wel bewijs was voor een verlenging na tamoxifen, maar nog niet na een aromataseremmer. Ook wordt er een overzicht gegeven van alle huidige studies, die vooral de waarde van verlengde hormoontherapie na aromataseremmers onderzoeken.

Eén van die studies is de IDEAL-studie, waarvan de primaire resultaten beschreven staan in **hoofdstuk 3**. In de IDEAL studie zijn patiënten die eerder behandeld zijn met 5 jaar hormoontherapie (tamoxifen, AI of een combinatie van beiden), gerandomiseerd tussen 2.5 en 5 jaar extra letrozol (een AI). De primaire uitkomstmaat was ziektevrije overleving. Na meer dan 5 jaar follow-up, was er geen verschil in ziektevrije overleving tussen beide groepen. Ook was er geen verschil in totale overleving. Wel werd een verschil gezien op het ontstaan van nieuwe borstkankers, dit was meer dan de helft minder in de langer behandelde groep (hazard ratio 0.39). Echter, in absolute zin was het verschil maar 2% (3% in 2.5-jaar groep, en 1% in 5-jaar groep, en hiermee is er dus niet voldoende waarde van de langere hormoontherapie voor de gehele groep.

Ondertussen was er ook een andere studie gepubliceerd, de eveneens Nederlandse DATA trial. Hierin werden patiënten na 3 jaar tamoxifen, gerandomiseerd tussen 3 of 6 jaar anastrozol. In die studie, werd er aangetoond dat voor patiënten met uitzaaiingen in de lymfeklieren er wel baat is bij langere therapie, terwijl dit niet voor de hele groep geldt. Daarom hebben wij in **hoofdstuk 4** een uitgebreidere subgroepanalyse gedaan in de IDEAL studie, om te kijken naar het effect van langere behandeling bij patiënten die alleen zijn voorbehandeld met tamoxifen én uitzaaiingen in de lymfeklieren hebben. Inderdaad blijkt in deze analyse dat er voor deze specifieke subgroep wél een voordeel is van langere therapie. Echter, de analyse is gebaseerd op een klein aantal patiënten, en er kunnen nog geen definitieve conclusies uit worden getrokken.

Een ander probleem bij verlengde hormoontherapie, is het feit dat veel patiënten niet in staat zijn om de behandeling vol te houden. Dit fenomeen hebben we onderzocht in **hoofdstuk 5.** In die IDEAL studie werd het merendeel van de uitval (58%) verklaard door bijwerkingen van de hormoontherapie. Vooral bijwerkingen als depressie, gewrichtspijn en vermoeidheid leidden vaak tot het stoppen van de behandeling. Opmerkelijk genoeg had het stoppen van hormoontherapie nadat er een bijwerking is opgetreden, geen nadelig effect op de overleving.

Biomarkers voor hormoontherapie

Eén van de manieren waarop patiënten beter geselecteerd kunnen worden voor de hormoontherapie, is het gebruik van biomarkers. Zoals hierboven reeds beschreven, wordt op dit moment in de tumor gekeken bij hoeveel procent van de tumorcellen de oestrogeenreceptor aanwezig is. Als dit bij meer dan 10% van de cellen is, noemen we de tumor hormoongevoelig. Echter, het feit dat deze receptor aanwezig is, wil niet per se zeggen dat deze ook actief is. In **hoofdstuk 6** beschrijven we een nieuwe en nog ongebruikte methode, waarmee we kunnen bepalen of de oestrogeenreceptor ook daadwerkelijk actief is. Hiermee hebben we gekeken naar patiënten die neo-adjuvant zijn behandeld met letrozol. Zowel voor als na de behandeling hebben we de activiteit van de oestrogeenreceptor bepaald. We hebben gezien dat bij patiënten die vooraf een hogere activiteit hadden, de behandeling beter aansloeg, en we hebben ook gezien dat bij deze patiënten de activiteit sterker af nam. Deze techniek kan wellicht in de toekomst gebruikt worden om patiënten te selecteren die meer, of juist geen baat hebben bij hormoontherapie. Als je in staat zou zijn om van tevoren te voorspellen wie er geen baat heeft, zou je de behandeling kunnen aanvullen met gerichte therapie zoals een CDK4/6 remmer, of vervangen door chemotherapie.[7]

Een andere mogelijke biomarker is het immuunsysteem, en tumorinfiltrerende T-cellen (TILs) in het bijzonder. Al eerder is aangetoond dat bij hormoongevoelig borstkanker, er geen prognostische waarde is van TILs, terwijl deze wel voorspellen voor een gunstige uitkomst bij hormoonongevoelig (met name triple-negatief) borstkanker. Echter, een biomarker kan predictieve waarde hebben, zonder op zichzelf prognostisch te zijn. Derhalve hebben wij in **hoofdstuk 7** onderzocht wat de predictieve waarde is van CD8-positieve (cytotoxische) TILs bij hormoongevoelig borstkanker. Dit onderzoek is gedaan in het Nederlandse deel van de IES studie, waarin patiënten werden gerandomiseerd tussen 5 jaar tamoxifen, of 2.5 jaar tamoxifen gevolgd door exemestaan (een AI). Onze resultaten tonen dat wanneer er weinig TILs aanwezig zijn, er veel voordeel is van de behandeling met exemestaan (HR 0.27, p<0.001), terwijl als er veel TILs zijn er geen verschil is tussen beide behandelingen (HR 1.34, p=0.36). Deze analyse is herhaald in de TEAM, waar in de eerste periode ook tamoxifen met exemestaan is vergeleken. Ook hierin werd gezien dat bij weinig TILs er een voordeel is van exemestaan (HR 0.67, p=0.048), maar niet bij veel TILs (HR 0.82, p=0.32), alhoewel het verschil in deze studie minder groot is. Als deze resultaten worden herhaald in een andere studie, zou dit kunnen leiden tot een beslissing over het type hormoontherapie op basis van de aanwezigheid van TILs.

Om te onderzoeken waarom TILs geen prognostische waarde hebben bij HR-positief, maar wel bij HR-negatieve tumoren, hebben we in **hoofdstuk 8** onderzocht of FAS hierin wellicht een rol speelt. FAS is beter bekend als de 'death receptor', en moet op een cel zitten om door een cytotoxische T-cel te kunnen worden aangevallen. Downregulatie van deze receptor kan dus zorgen voor een escape van het immuunsysteem, en zou een strategie van de tumor kunnen zijn. We hebben gezien dat FAS twee keer zoveel tot expressie komt in HR-negatieve tumoren. Daarnaast hebben we gezien dat CD8-positieve TILs bij HR-negatieve tumoren, alleen maar prognostisch zijn als ook FAS tot expressie komt in deze tumoren. De combinatie van deze twee conclusies zou wellicht (deels) kunnen verklaren waarom TILs wel prognostisch zijn bij HR-negatieve tumoren.

Een andere manier om biomarkers te gebruiken bij borstkanker, is het gebruik van genexpressie. Er zijn meerdere commerciële tests beschikbaar, die op basis van de expressie van verschillende genen in de tumor het risico op terugkeer van de ziekte kunnen bepalen. Hiermee kan dus vooraf een inschatting worden gemaakt of de patiënt een hoog of laag risico heeft op terugkeer van de ziekte, waardoor je in theorie meer of minder winst van de behandeling kunt verwachten. In **hoofdstuk 9** presenteren we een uitgebreide systematische review over de vier meest gebruikte testen in Europa: OncotypeDX, MammaPrint, Endopredict en Prosigna. Deze review richt zich op 4 belangrijke aspecten: de ontwikkeling van de test, de klinische validatie, het effect van de testen op gebruik van chemotherapie, en de economische waarde van de test. Wij concluderen dat zowel de Amerikaanse OncotypeDX als de Nederlandse MammaPrint goed zijn onderzocht, en op enkele verschillen na gelijkwaardig zijn voor klinisch gebruik. Voor beide testen zijn al grote klinische studies verricht om de waarde in de praktijk te laten zien. Voor de MammaPrint is dat de MINDACT studie geweest, waarover we in **hoofdstuk 10** een ingezonden brief schrijven waarin we oproepen om goed te kijken bij welke subgroepen een dergelijke gentest echt meerwaarde heeft, en bij welke subgroep een bepaalde risico-uitslag al voor de hand ligt vanwege de klinische karakteristieken. Indien genexpressie testen daadwerkelijk bij de juiste subgroep ingezet kunnen gaan worden, zullen deze wellicht ook een rol kunnen spelen bij de selectie van patiënten met zo'n gunstig risicoprofiel, dat hormoontherapie voor deze patiënten weinig toegevoegde waarde zal gaan hebben.

Samenvattend behandelt dit proefschrift diverse aspecten van personalisatiemogelijkheden voor de hormoontherapie van patiënten met hormoongevoelig

borstkanker, zowel qua behandelduur als qua inzet van diverse soorten biomarkers. In de toekomst zal deze personalisatie over het algemeen waarschijnlijk leiden tot een gerichter gebruik van hormoontherapie, omdat we beter in staat zullen zijn om de patiënten te selecteren die baat hebben bij de therapie, in plaats van dat vrijwel alle patiënten vrijwel ongericht behandeld worden. Daarnaast zal er een klein deel zijn waarin de hormoontherapie juist geëscaleerd wordt tot een langere therapieduur of een gebruik van aromataseremmers in plaats van tamoxifen. Hiermee zal de behandeling van de patiënt met borstkanker, en daarmee ook de prognose, nog verder kunnen verbeteren.

References

1. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
2. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-1717.
3. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. Lancet 2013;381:805-816.
4. Gray R, Rea D, Handley K. aTTom: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years in 6,953 women with early breast cancer. J Clin Oncol 2013;31(supplements; abstract 5).
5. Goss PE, Ingle JN, Martino S et al. A Randomized Trial of Letrozole in Postmenopausal Women after Five Years of Tamoxifen Therapy for Early-Stage Breast Cancer. N Engl J Med 2003;349(19):1793-1802.
6. Al-Mubarak M, Tibau A, Templeton AJ et al. Extended Adjuvant Tamoxifen for Early Breast Cancer: A Meta-Analysis. PLoS ONE 2014;9(2):e88238.
7. de Groot AF, Kuijpers CJ, Kroep JR. CDK4/6 inhibition in early and metastatic breast cancer: A review. Cancer Treatment Reviews 2017;60(Supplement C):130-138.

List of publications

2013

Blok EJ, Kuppen PJK, van Leeuwen JEM et al. Cytoplasmic overexpression of HER2: a key factor in colorectal cancer? Clinical Medicine Insights: Oncology 2013;7:S10811.

2015

Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.

2016

Blok EJ, van de Velde CJH, Smit VTHBM. Letter to the editor: 70-Gene Signature in Early-Stage Breast Cancer. New England Journal of Medicine 2016;375(22):2199

2017

Blok EJ, van den Bulk J, Dekker-Ensink NG et al. Combined evaluation of the FAS cell surface death receptor and CD8+ tumor infiltrating lymphocytes as a prognostic biomarker in breast cancer. Oncotarget 2017; 8(9): 15610.

Derks MGM*, **Blok EJ***, Seynaeve C et al. Adjuvant tamoxifen and exemestane in women with postmenopausal early breast cancer (TEAM): 10-year follow-up of a multicentre, open-label, randomised, phase 3 trial. The Lancet Oncology 2017;18(9):1211-1220.

2018

Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg E et al. Optimal Duration of Extended Adjuvant Endocrine Therapy for Early Breast Cancer; Results of the IDEAL Trial (BOOG 2006-05). Journal of the National Cancer Institute 2018;110(1):djx134.

Blok EJ, Bastiaannet E, van den Hout WB et al. Systematic review of the clinical and economic value of gene expression profiles for invasive early breast cancer available in Europe. Cancer Treatment Reviews 2018;62:74-90

Swets M, Kuppen PJK, **Blok EJ** et al. Are pathological high-risk features in locally advanced rectal cancer a useful selection tool for adjuvant chemotherapy? European Journal of Cancer 2018;89:1-8

Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg WM et al. Relevant factors for the optimal duration of extended endocrine therapy in early breast cancer. Breast Cancer Research and Treatment 2018;168(2):413-420

Blok EJ, Kroep JR, Meershoek-Klein Kranenbarg WM et al. Treatment decisions and the impact of adverse events before and during extended endocrine therapy in postmenopausal early breast cancer. European Journal of Cancer 2018; *in press*

Blok EJ, Engels CC, Dekker-Ensink NG et al. Exploration of tumour-infiltrating lymphocytes as a predicitve biomarker for adjuvant endocrine therapy in early breast cancer. Breast Cancer Research and Treatment 2018; *accepted for publication*

*Shared first authors

List of co-authors

A. van Brussel
A. van de Stolpe
A.C. Charehbili
C. Kanters
C.C. Engels
C.J.H. van de Velde
C.S. Seynaeve
E. Bastiaannet
E. den Biezen-Timmermans
E.J.Th. Rutgers
G.J. Liefers
H. Putter
J. van den Bulk
J.J.M. van der Hoeven
J.M. Bliss
J.M.S. Bartlett
J.P. Morden
J.R. Kroep
J.W.R. Nortier
M. Alves de Inda
M. Duijm-de Carpentier
M.G.M. Derks
N.G. Dekker-Ensink
P.J.K. Kuppen
R. Derr
R.C. Coombes
S. Fruytier
S.C. Linn
V.T.H.B.M. Smit
W. Verhaegh
W.B. van den Hout
W.M. Meershoek-Klein Kranenbarg

Curriculum Vitae

Erik Jan Blok was born on March 6th 1989 in Papendrecht, where he grew up with his parents and younger sister Ilse. He graduated from the Johan de Witt Gymnasium in 2007, and started with Biomedical Sciences at Leiden University.

After obtaining his Bachelor's degree in Biomedical Sciences in 2010, he was accepted for the 'Dubbeltraject', a special program for selected biomedical students to start with medical school. In 2011, he also obtained his Bachelor's degree in Medicine and started with two Master's programs, combining both Medicine and Biomedical Sciences. He finished both Master's programs in 2014 and 2015 respectively, of which the latter with the addition *cum laude*. During his study years, he participated as a treasurer in the organizing committee of the Leiden International Medical Students Conference (LIMSC 2011), one of the largest medical students conferences in the world, and in multiple other committees.

His first experience with cancer research was during his Biomedical Sciences Bachelor's internship, in which he evaluated the use of anti-carbohydrate autoantibodies as a screening tool for colorectal cancer at the Department of Parasitology, supervised by dr. M. Wuhrer. During his Master's internships, he evaluated HER2 as a target for image-guided surgery at the Department of Surgery, supervised by dr. C.F.M. Sier and dr. P.J.K. Kuppen, and also assessed predictive and prognostic biomarkers for the response to (neo)adjuvant endocrine therapy in early breast cancer, supervised by dr. J.R. Kroep and prof. dr. C.J.H. van de Velde.

His final internship was the starting point for his PhD program and the current thesis, in which he combined translational and clinical research with the focus on endocrine therapy in early breast cancer, supervised by dr. P.J.K. Kuppen, dr. J.R. Kroep and prof. dr. C.J.H. van de Velde. In January 2018, he started working as a non-training resident in Internal Medicine in the Reinier de Graaf hospital in Delft, with the intention to apply for a training position in Internal Medicine later that year.

Dankwoord

Prof. dr. Vd Velde, beste professor, ik wil u heel hartelijk danken dat ik de afgelopen jaren deel heb mogen uitmaken van uw onderzoeksgroep. Ik vond het een grote eer met een internationale toponderzoeker als u te mogen samenwerken. Uw wijze lessen over politiek in de wetenschap zullen mij altijd bij blijven, en hebben mij tot een betere onderzoeker gemaakt.

Dr. Kroep, beste Judith. Heel erg bedankt voor de mooie jaren als promovendus. Jouw persoonlijke benadering, in combinatie met medische kennis van zaken, werklust en vertrouwen, maken van jou een fijne begeleider. Heel veel dank voor al het gestelde vertrouwen.

Dr. Kuppen, beste Peter. Al vanaf mijn studententijd heb jij me onder je hoede genomen, en hebben we veel samen gewerkt. Je tomeloze inzet op verschillende vlakken, zowel op tumorimmunologisch gebied als ook in de begeleiding van studenten is bewonderenswaardig. Dank voor al je vertrouwen!

Prof. dr. Smit, beste Vincent. De urenlange sessies achter de microscoop hadden mij bijna overtuigd om voor de pathologie te kiezen, een prachtig vak midden in het web van de oncologie. Heel veel dank voor alle hulp en inspiratie.

Dr. Liefers, beste GJ. Jouw frisse blik en charisma is een ware inspiratiebron voor velen. Je hebt mij laten zien hoe je door enthousiasme, slimheid en de juiste dosis nonchalance veel voor elkaar kunt krijgen. Dank voor deze wijze lessen!

Beste collega's van het datacentrum Heelkunde, en in het bijzonder Elma, jullie werk in het oncologisch onderzoek is onmisbaar. De kwaliteit van jullie databases, en ook de jarenlange ervaring met onderzoek, hebben mij veel werk en zorgen uit handen genomen, waarvoor hartelijk dank.

Beste collega's van het Heelkunde Research lab, zonder jullie was dit proefschrift een stuk leger geweest. Jullie jarenlange kennis en expertise op het gebied van translationeel onderzoek heeft mij enorm geholpen, dank hiervoor!

Dear international co-authors, thank you for showing me the world of international collaborations in medical research. These were important lessons I will remember for the rest of my career.

Beste collega's van de Medische Oncologie, en in het bijzonder Florine en Stefanie. Hoewel onze projecten niet altijd even succesvol waren, was de samenwerking dat zeker wel. Ook waren jullie mijn vaste reisgenootjes, van luxe vergaderingen in Bordeaux tot legendarische etentjes in San Antonio. Heel veel dank hiervoor.

Lieve collega-onderzoekers van de Heelkunde; lunchmaatjes, fietsvrienden, reisgenoten, voetbalbuddies, wat hebben we een leuke tijd met elkaar gehad. Je realiseert je niet dat de collega's om je heen zo'n grote rol spelen in je leven, totdat je momenten van tegenslag tegenkomt. Ik wil jullie enorm bedanken voor de gezelligheid, inspirerende discussies, en steun in moeilijkere tijden. Ik ben heel blij dat ik veel van jullie inmiddels tot mijn vriendenkring mag rekenen.

Beste kamergenoten, wat heb ik door jullie een leuke tijd gehad als onderzoeker op de J10-71. We deelden letterlijk lief en leed met elkaar, en door jullie kwam ik iedere dag met veel plezier naar het LUMC. In het bijzonder Annemarie en Graziella, jullie rol als kamermentor is onvervangbaar. Martine, Nikki, Noor en Yvette, jullie waren er voor me in de periode dat ik dat het hardste nodig had. Dit zal ik nooit vergeten.

Dubbeltraject-vrienden, in 2009 begonnen wij samen aan een reis, waarvan we niet wisten waar die zou eindigen. Inmiddels hebben we allemaal onze passie gevonden, en ik vond het een eer om dat proces met jullie, de crème de la crème van de Leidse BW-studenten, te mogen meemaken.

Yotam en Gido, ik ben enorm trots dat ik jullie tot mijn vrienden mag rekenen. Jullie zijn niet alleen de meest intelligente, maar ook de meest sociale mensen die ik ken. De soms messcherpe academische discussies hebben een wezenlijke bijdrage geleverd aan dit proefschrift. Ik wil jullie daarvoor hartelijk bedanken.

Opa en oma, jullie hebben mij de definitie van een liefdevolle familie geleerd. Ik ben heel blij dat jullie er nog zijn, en ondanks de vele tegenslagen nog fier in het leven staan. Opa, ondanks de ogenschijnlijke lichamelijke beperkingen, bent u een van de slimste en meest inspirerende mensen die ik ken. Dank voor alle wijze lessen.

Mama en Ilse, de vele mini-vakanties op de Elzenzoom zorgden naast een lege wasmand en een gevulde maag ook altijd voor een weer opgeladen batterij. Wat zijn jullie sterk geweest de afgelopen jaren, en hoe goed hebben jullie de draad van het leven weer opgepakt. Heel veel dank voor jullie steun, trots en vertrouwen.

Pap, hoewel je dit proefschrift nooit zal kunnen vasthouden, weet ik dat je trots op me bent.